GENOME-SCALE MODELING OF REDOX METABOLISM AND THERAPEUTIC RESPONSE IN RADIATION-RESISTANT TUMORS

A Dissertation
Presented to
The Academic Faculty

by

Joshua Elliott Lewis

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Bioinformatics in the
Wallace H. Coulter Department of Biomedical Engineering

Georgia Institute of Technology & Emory University
May 2020

GENOME-SCALE MODELING OF REDOX METABOLISM AND THERAPEUTIC RESPONSE IN RADIATION-RESISTANT TUMORS

Approved by:

Dr. Melissa L. Kemp, PhD, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. David S. Yu, MD, PhD
Department of Radiation Oncology
Emory University

Dr. Greg Gibson, PhD
School of Biological Sciences
Georgia Institute of Technology

Dr. Lee A. Cooper, PhD
Department of Pathology
Northwestern University

Dr. Eberhard O. Voit, PhD
Department of Biomedical Engineering
Georgia Institute of Technology

Date Approved: March 20, 2020

To my family

ACKNOWLEDGEMENTS

My development as a scientist can be attributed to the synergistic involvement of a number of people who I would like to gratefully acknowledge. First, I would like to thank my PhD advisor, Dr. Melissa Kemp, for providing me with tremendous guidance, challenging me to continually improve, and instilling in me a passion for scientific research. Next, I would like to thank my committee members for their valuable input and advice throughout my PhD training: Dr. Lee Cooper, Dr. Greg Gibson, Dr. Eberhard Voit, and Dr. David Yu. I would also like to thank research collaborators Dr. Cristina Furdui and Dr. David Boothman, as well as my previous and current financial support through the MD/PhD program (NIH T32 GM008169-29), Computational Biology and Predictive Health Training Grant (NIH T32 GM105490), and Ruth L. Kirschstein National Research Service Award (NIH F30 CA224968).

I would like to thank the members of the Kemp lab, both past and present, for providing an enjoyable and intellectually stimulating environment to work and learn, as well as bagels during morning lab meetings. I would also like to thank my friends, especially those in the Emory MD/PhD program, who have made this long and arduous journey so much more bearable.

Finally, I would like to thank my family for their unwavering support and acceptance that I will still be in school for a very long time. I especially want to thank my parents, who made countless sacrifices to allow me to pursue my dreams and try to make a positive difference in the world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

[E]	Enzyme abundance
•OH	Hydroxyl radical
10-formyl-THF	10-formyltetrahydrofolate
12HpETE	12-hydroxyperoxyeicosatetraenoic acid
13BPG	1,3-bisphosphoglycerate
2HG	2-hydroxyglutarate
2PG	2-phosphoglycerate
3HB	3-hydroxybutyrate
3HBCoA	3-hydroxybutyrl-CoA
3PG	3-phosphoglycerate
3PG	3-phosphoglycerate
4K-CPA	4-Ketocyclophosphamide
4OH-CPA	4-Hydroxycyclophosphamide
5-MTHF	5-methyltetrahydrofolate
5,10-CH	THF = 5,10-methenyltetrahydrofolate
5,10-CH_2-THF	5,10-methylenetetrahydrofolate
6PG	6-phosphogluconate
6PGL	6-phosphogluconolactone
αKG	Alpha-ketoglutarate
AA	Acetoacetate
AA	Arachidonic acid
AACoA	Acetoacetyl-CoA

ACC	Adrenocortical carcinoma
ACoA	Acetyl-CoA
ACR	Acrolein
ACR-GSH	Acrolein-GSH
ALCPh	Alcophosphamide
ALDH	Aldehyde dehydrogenase
ALDPh	Aldophosphamide
AMP	Adenosine monophosphate
AMPD	Adenosine monophosphate deaminase
Arg	Arginine
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
AUROC	Area under the receiver operating characteristic curve
BLCA	Bladder urothelial carcinoma
BRCA	Breast invasive carcinoma; Breast cancer gene
CAA	Chloroacetaldehyde
CAT	Catalase
CAT	Catalase
CBR	Carbonyl reductase
CCLE	Cancer Cell Line Encyclopedia
CDP	Cytosine diphosphate
CDP-DAG	CDP-diacylglycerol
CESC	Cervical and endocervical cancers
Cis	Cisplatin
Cis-GSH	Cisplatin-GSH

Cis-MT	Cisplatin-metallothionein
Cis-mtDNA	Cisplatin-mitochondrial DNA
Cit	Citrate
CL	Cardiolipin
COAD	Colon adenocarcinoma
CPA	Cyclophosphamide
CPh	Carboxyphosphamide
dATP	Deoxyadenosine triphosphate
Daun	Daunosamine
DCE-CPA	Dechloroethylcyclophosphamide
DG	Diacylglycerol
dGDP	Deoxyguanosine diphosphate
DGLA	Dihomo-γ-linolenic acid
dGMP	Deoxyguanosine monophosphate
dGTP	Deoxyguanosine triphosphate
DHAP	Dihydroxyacetone phosphate
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DLBC	Diffuse large B-cell lymphoma
DMEM	Dulbecco's Modified Eagle's Medium
DMEM/F-12	1:1 Dulbecco's Modified Eagle's Medium:Ham's F-12 Medium
Dox	Doxorubicin
Dox-deoxy	7-Deoxydoxorubicinone
Dox-ol	Doxorubicinol
Dox-Q	Doxorubicin, quinone

Dox-SQ	Doxorubicin, semiquinone
EC	Enzyme commission
EC	Enzyme commission
E_{hc}	Half-cell potential
F16BP	Fructose 1,6-bisphosphate
F1P	Fructose 1-phosphate
F26BP	Fructose 2,6-bisphospahte
F6P	Fructose 6-phosphate
FA-CoA	Fatty acyl-CoA
FAD	Flavin adenine dinucleotide
$FADH_2$	Flavin adenine dinucleotide, reduced
FBA	Flux balance analysis
FBS	Fetal bovine serum
FFA	Free fatty acid
Fru	Fructose
Fuc	Fucose
Fum	Fumarate
FVA	Flux variance analysis
G3P	Glyceraldehyde 3-phosphate
G6P	Glucose 6-phosphate
G6PD	Glucose 6-phosphate dehydrogenase
GBM	Gradient boosting machine; glioblastoma multiforme
GDC	Genomic Data Commons
GDISC	Gene-Drug Interaction for Survival in Cancer
GDP	Guanosine diphosphate

GDP	Guanosine diphosphate
GDP-ddM	GDP-4-keto-6-deoxymannose
GDP-Fuc	GDP-fucose
GDP-M	GDP-mannose
GLA	γ-linolenic acid
Glc	Glucose
GLUD	Glutamate dehydrogenase
Glyald	Glyceraldehyde
Glyc3P	Glycerol 3-phosphate
GMP	Guanosine monophosphate
GPR	Gene-protein-reaction rule
GPx	Glutathione peroxidase
Grx	Glutaredoxin
GSH	Glutathione, reduced
GSR	Glutathione reductase
GSSG	Glutathione disulfide
GTP	Guanosine triphosphate
Gylc	Glycerol
H_2O_2	Hydrogen peroxide
HILIC	Hydrophilic interaction liquid chromatography
HMGCoA	3-hydroxy-3-methylglutaryl-CoA
HNSC	Head and neck cancer
HNSCC	Head and neck squamous cell carcinoma
HPV	Human papillomavirus
iBMK	Immortalized baby mouse kidney epithelial

ICit	Isocitrate
IDH	Isocitrate dehydrogenase
IMP	Inosine monophosphate
IMPDH	Inosine monophosphate dehydrogenase
k_{cat}	Catalytic rate
KO	Knockout
LA	Linoleic acid
Lac	Lactate
LGG	Lower glade glioma
LIHC	Liver hepatocellular carcinoma
LPA	Lysophosphatidic acid
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
M16BP	Mannose 1,6-bisphosphate
M1P	Mannose 1-phosphate
M6P	Mannose 6-phosphate
MAE	Mean absolute error
MAG	Monoacylglycerol
Mal	Malate
MCoA	Malonyl-CoA
ME	Malic enzyme
MSE	Mean squared error
MTHFD	Methylenetetrahydrofolate dehydrogenase
MTHFR	Methylenetetrahydrofolate reductase
NAAD	Nicotinic acid dinucleotide

NAD^+	Nicotinamide adenine dinucleotide, oxidized
NADH	Nicotinamide adenine dinucleotide, reduced
NADK	NAD kinase
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidized
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
NADS	NAD synthase
NAM	Nicotinamide
NAMN	Nicotinic acid mononucleotide
NAMPT	Nicotinamide phosphoribosyltransferase
NBCI	National Center for Biotechnology Information
NMN	Nicotinamide mononucleotide
NMNAT	Nicotinamide mononucleotide adenylyltransferase
NNT	Nicotinamide nucleotide transhydrogenase
NOX	NADPH oxidase
NPV	Negative predictive value
NQO1	NAD(P)H:quinone oxidoreductase 1
NR	Nicotinamide riboside
Nrf2	Nuclear factor (erythroid-derived 2)-like 2
NRK	Nicotinamide riboside kinase
NT5	5'-nucleotidase
$O_2^{\bullet-}$	Superoxide anion
OAA	Oxaloacetate
OAA	Oxaloacetate
PA	Phosphatidic acid
PARP	Poly(ADP-ribose)polymerase

PC	Phosphatidylcholine
PCoA	Palmitoyl-CoA
PDC	Pyruvate dehydrogenase complex
PE	Phosphatidylethanolamine
PEP	Phosphoenolpyruvate
PG	Phosphatidylglycerol
PGD	Phosphogluconate dehydrogenase
PGG2	Prostaglandin G2
PGP	Phosphatidylglycerol-phosphate
PhMust	Phosphoramide mustard
PhMust-GSH	Phosphoramide mustard-GSH
PI	Phosphatidylinositol
PNP	Purine-nucleotide phosphorylase
PPM	Phosphopentomutase
PPV	Positive predictive value
PRAD	Prostate adenocarcinoma
PRPP	Phosphoribosyl pyrophosphate
PRPS	Phosphoribosyl pyrophosphate synthase
Prx	Peroxiredoxin
PS	Phosphatidylserine
Pyr	Pyruvate
QA	Quinolinic acid
QPRT	Quinolinate phosphoribosyltransferase
R1P	Ribose 1-phosphate
R5P	Ribose 5-phosphate

READ	Rectum adenocarcinoma
RECIST	Response evaluation criteria in solid tumors
ROC	Receiver operating characteristic
ROS	Reactive oxygen species
S.E.	Standard error
SHAP	Shapley additive explanations
SKCM	Skin cutaneous melanoma
SOD	Superoxide dismutase
STAD	Stomach adenocarcinoma
Suc	Succinate
SucCoA	Succinyl-CoA
TCA	Tricarboxylic acid cycle
TCGA	The Cancer Genome Atlas
TG	Triglyceride
THCA	Thyroid carcinoma
THF	Tetrahydrofolate
TR	Thioredoxin reductase
Trp	Tryptophan
Trx	Thioredoxin
TXA2	Thromboxane A2
UCEC	Uterine corpus endometrial carcinoma
UCS	Uterine carcinosarcoma
V_{max}	Maximum flux
WT	Wild-type
XDH	Xanthine dehydrogenase/xanthine oxidase

XMP	Xanthosine monophosphate
ΔG	Change in Gibbs free energy
ns	not significant
*	$p \leq 0.05$
**	$p \leq 0.01$
***	$p \leq 0.001$
****	$p \leq 0.0001$

SUMMARY

Despite being one of the oldest forms of cancer therapy and still a primary treatment modality, radiation therapy is not effective across all cancer types and tumor resistance to radiation is still not well understood. As our ability to characterize tumor pathophysiology increases with new -omic technologies, a broad clinical goal is prognostic indicators of therapeutic outcomes for personalizing therapeutic regimens. While redox metabolism is a known factor, methods for analyzing systems-level involvement of cellular metabolism in radiation response have not been previously developed. This dissertation presents the construction of novel genome-scale Flux Balance Analysis (FBA) models of individual radiation-sensitive and -resistant patient tumors from The Cancer Genome Atlas (TCGA) to explore the role of redox metabolism in radiation sensitivity, to identify diagnostic and therapeutic biomarkers for radiation response, and to predict response to radiation-sensitizing chemotherapies in radiation-resistant tumors.

A novel bioinformatics platform was developed to integrate genomic, transcriptomic, kinetic, and thermodynamic parameters from 716 radiation-sensitive and 199 radiation-resistant TCGA tumors into personalized genome-scale FBA models. Pan-cancer model predictions identified increased mitochondrial production of redox cofactors, including NADPH and glutathione, as well as increased H_2O_2-scavenging fluxes in radiation-resistant tumors. Simulated gene knockout screens were utilized to discover novel targets in redox metabolism, central carbon metabolism, and folate metabolism which differentially impact antioxidant production and ROS clearance in radiation-resistant tumors; these targets were experimentally validated through siRNA gene

knockdown in matched radiation-sensitive and -resistant cancer cell lines among multiple cancer types. Finally, personalized metabolic flux profiles were generated for individual radiation-resistant cancer patients to identify optimal targets for radiation sensitization. This work not only improved upon methodological shortcomings of previous FBA models of cancer metabolism, but is the first to utilize genome-scale modeling for identifying metabolic differences between radiation-sensitive and -resistant tumors that could be exploited for improving radiation sensitivity.

Machine learning classifiers were developed which integrate multi-omic data from TCGA patients and novel metabolic outcomes from personalized FBA models to predict radiation sensitivity. A dataset-independent ensemble architecture with gradient boosting models and Bayesian optimization yielded improved predictive accuracy and biomarker detection compared to previously-developed classifiers for radiation response. Experimentally-validated predictions of metabolite production from radiation-sensitive and -resistant FBA tumor models were integrated into multi-omic classifiers; metabolites involved in lipid metabolism, nucleotide metabolism, and immune modulation were identified as having significant associations with radiation response. Subgroups of patients with differing utilities of clinical versus metabolomic datasets for radiation response prediction were discovered, and personalized panels of multi-omic and non-invasive biomarkers with optimal diagnostic utility were developed. This work made significant advancements by being the first to integrate FBA model predictions into machine learning classifiers for cancer treatment outcomes.

Finally, FBA models of radiation-resistant TCGA tumors were used to predict response to radiation-sensitizing chemotherapies and investigate their effects on tumor

redox metabolism. A novel multi-feature FBA objective function screen was developed, resulting in significant improvements in model predictions of treatment response, as well as identification of redox cofactors directly involved in drug metabolism. The radiation-sensitizing effect of chemotherapeutic treatment was predicted in radiation-resistant tumors by assessing drug-associated decreases in antioxidant levels, and machine learning regressors were utilized to identify multi-omic biomarkers from patient tumors which are associated with increased radiation sensitization. This work was the first to utilize genome-scale modeling to assess the role of chemotherapeutic treatment on tumor redox metabolism and radiation sensitization.

In summary, a generalizable framework for creating genome-scale metabolic models of individual patient tumors was developed. The collective properties of these personalized models improved pathophysiological insights into the role of redox metabolism in the tumor responses to radiation and radiation-sensitizing chemotherapies. This framework resulted in a reduced set of clinically-useful biomarkers for both the *a priori* prediction of radiation response as well as targeted sensitization of radiation-resistant tumors to radiation therapy. This personalized medicine approach represents a paradigm shift in how diagnostic and treatment strategies for radiation-resistant cancer patients are developed, ultimately improving the standard of care for these patients.

CHAPTER 1. INTRODUCTION

Tumor resistance to radiation therapy remains a significant obstacle to long-term cancer patient survival [1, 2]. As multi-omic characterizations of patient tumors are currently being developed, personalized biomarkers with diagnostic and therapeutic utility in the management of radiation-resistant cancer patients are under investigation [3]. Redox metabolism, involving the production of reducing cofactors and scavenging of reactive oxygen species (ROS), has been previously implicated in the mechanism of action of ionizing radiation therapy as well as the cellular response to both radiation treatment and radiation-sensitizing chemotherapies [4-6]. However, because redox cofactors such as NADPH, NADH, and GSH are involved in hundreds of interconnected reactions throughout the human metabolic network, experimental studies focusing on individual metabolic targets have not yielded much success for improving radiation sensitivity in patient tumors [7]. Additionally, previous approaches to biomarker detection which utilized only single -omics datatypes lack the predictive accuracy and applicability necessary for clinical use [8]. Instead, computational approaches which integrate multi-omic tumor data with genome-scale metabolic networks are needed to model redox metabolism and its involvement in the tumor response to radiation therapy.

Flux balance analysis (FBA) is a metabolic modeling approach for integrating biological information from samples of interest to predict steady-state reaction fluxes throughout the human metabolic network [9]. While FBA has been used previously to model cancer metabolism, personalized FBA models of patient tumors have not been previously developed to study the role of redox metabolism in radiation sensitivity [10].

Also, FBA models have yet to be integrated with machine learning classifiers for the prediction of novel diagnostic biomarkers, nor have they been used to model tumor chemotherapy response and radiation sensitization. Methodological shortcomings in FBA model development, including non-physiological utilization of gene expression data and failure to integrate kinetic and thermodynamic constraints, have limited the clinical utility of previous modeling attempts towards identifying diagnostic and therapeutic targets in cancer [11, 12].

1.1 Research Objectives and Specific Aims

The *objective* of this project was to develop a novel FBA-based bioinformatics platform for integrating multi-omic tumor data from The Cancer Genome Atlas (TCGA) into personalized models of radiation-sensitive and radiation-resistant tumors. These models would be used to 1) identify novel targets of redox metabolism in radiation-resistant tumors, 2) integrate metabolic information from patient tumors into machine learning models for prediction of radiation therapy response, and 3) model radiation-sensitizing chemotherapies to identify in which radiation-resistant patient tumors these therapies would be optimal alone or in conjunction with radiation treatment. The *outcomes* of this project are 1) a generalizable framework for developing accurate genome-scale metabolic models of individual patient tumors, 2) improved pathophysiological insights into the role of redox metabolism in the tumor response to radiation and radiation-sensitizing chemotherapies, and 3) identification of a reduced set of clinically-useful biomarkers for both the *a priori* prediction of radiation response as well as targeted sensitization of

radiation-resistant tumors to radiation therapy. This objective was addressed by the following three aims:

Aim 1: Generate personalized FBA metabolic models of patient tumors using multi-omic data from TCGA and compare differences in redox metabolism between radiation-sensitive and radiation-resistant tumors. The *technical hypothesis* was that methodological improvements to FBA model development would allow for accurate predictions of redox metabolism in patient tumors that could be experimentally validated in radiation-sensitive and -resistant cancer cell lines. The *biological hypothesis* was that radiation-resistant tumors are capable of upregulating production of redox cofactors for antioxidant defense against ROS-generating ionizing radiation, and that the enzymes involved in cofactor production are essential for viability of radiation-resistant cancers. Chapter 3 details early development of FBA models for matched radiation-sensitive (SCC-61) and radiation-resistant (rSCC-61) HNSCC cell lines through the novel integration of transcriptomic, kinetic, and thermodynamic data from these individual cell lines into maximum reaction flux constraints. These models demonstrated re-routing of NADPH-producing metabolic flux in rSCC-61 and accurately predicted gene targets which selectively impacted NADPH-driven sensitization to the redox-cycling chemotherapeutic ß-lapachone. Chapter 4 details the extension of this modeling framework for integrating multi-omic data into personalized FBA models of individual radiation-sensitive and -resistant TCGA patient tumors. Tumor-specific models were used to identify novel targets in the redox metabolic network which differentially impacted redox cofactor production and ROS-scavenging in radiation-resistant tumors; these targets were experimentally

validated through siRNA gene knockdown experiments in matched radiation-sensitive and -resistant cancer cell lines from multiple cancer types.

Aim 2: Develop machine learning classifiers for identification of novel multi-omic biomarkers for radiation therapy response. The *technical hypothesis* was that integration of metabolic predictions from FBA models of TCGA tumors into machine learning classifiers would yield novel multi-omic biomarkers that could be used for *a priori* prediction of radiation sensitivity. The *biological hypothesis* was that radiation-sensitive and -resistant tumors display differences in genomic, transcriptomic, and metabolomic features involved in redox biology and metabolism that enrich top predictive factors for machine learning classifiers of radiation response. Chapter 5 details the development of a dataset-independent ensemble architecture for integrating multi-omic datasets into machine learning models, yielding improved predictive accuracy and biomarker identification compared to previous radiation response classifiers. Additionally, FBA models were shown to accurately predict relative metabolite production between radiation-sensitive and -sensitive tumors, which was validated by experimental metabolomics in matched radiation-sensitive and -resistant cancer cell lines from multiple cancer types; these features were integrated into machine learning models to yield novel metabolic biomarkers for prediction of radiation response, including in non-invasive applications.

Aim 3: Incorporate the mechanism of action of radiation-sensitizing chemotherapies into FBA models to identify therapeutic strategies in individual radiation-resistant patients. The *technical hypothesis* was that utilization of a multi-feature objective function screen would significantly improve the accuracy of FBA model-predicted sensitivity to radiation-sensitizing chemotherapies. The *biological hypothesis*

was that tumor response to radiation-sensitizing chemotherapy relies on network-level metabolism of redox cofactors that are upregulated in radiation-resistant tumors. Chapter 6 details the integration of chemotherapy mechanisms of action into FBA models and identification of optimal FBA objective functions for prediction of chemotherapy response in radiation-resistant tumors. Additionally, possible mechanisms of redox-based radiation sensitization by these chemotherapies were identified and used to discover novel multi-omic biomarkers predictive for enhanced radiation-sensitizing effect in individual radiation-resistant tumors.

CHAPTER 2. BACKGROUND

2.1 Radiation Therapy

2.1.1 Epidemiology

Radiation therapy remains a cornerstone of cancer treatment; approximately 50% of all cancer patients receive radiation as part of their treatment regimen, and radiation therapy contributes an estimated 40% towards curative treatment in these patients [1, 13, 14]. However, tumor resistance to radiation therapy remains a significant obstacle to long-term cancer patient survival. Among cancer patients in The Cancer Genome Atlas (TCGA) database, 28.0% of patients did not show complete resolution of their tumor after treatment with radiation therapy [2]. Figure 2-1 shows that the proportion of radiation-resistant patients (RECIST criteria: stable disease or progressive disease) compared to radiation-sensitive patients (RECIST criteria: complete response or partial response) varies significantly between different cancer types; while more than 90% of TCGA patients with breast and thyroid cancers are sensitive to radiation therapy, over 70% of low-grade and high-grade glioma patients are radiation-resistant. Despite this observed variability in radiation response between cancer types, generalized pathophysiological principles of radiation resistance that apply across cancer types have been identified, and approaches for identifying novel biomarkers for improved diagnosis and treatment have included both cancer-specific and pan-cancer strategies [15, 16].

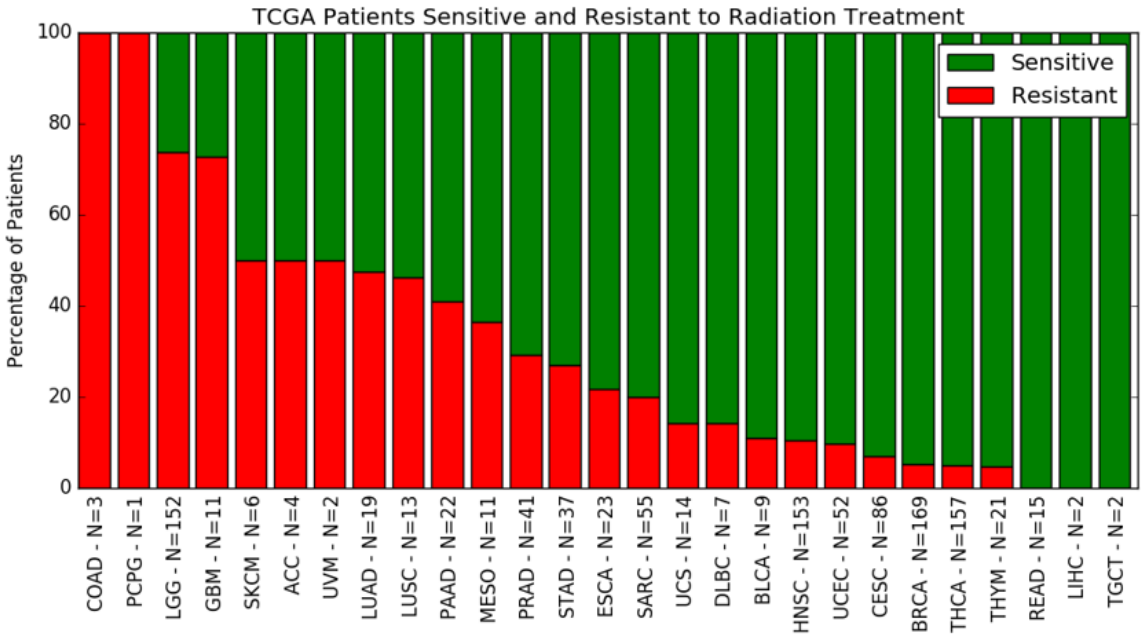


Figure 2-1. Comparison of number of radiation-sensitive and radiation-resistant TCGA patients between different cancer types.

2.1.2 Mechanism of action

Ionizing radiation from X-rays, gamma rays, or other sources kill cancer cells through damage to cellular structures, most notably DNA [17]. Radiation can act directly on DNA to cause damage, or can act indirectly through the production of reactive oxygen species (ROS) such as hydroxyl radical (•OH) and hydrogen peroxide (H_2O_2) derived from ionization of water. An estimated 60% of DNA damage from low-LET (linear energy transfer) radiations including X-rays and gamma rays is caused by indirect rather than direct effects [18]. Multiple types of DNA damage can occur; however, double-stranded breaks (DSB's) are considered the most lethal [19]. ROS cause single-stranded breaks (SSB's) through a Fenton reaction involving metal ions such as iron [4, 20]:

$$\begin{gathered} Fe^{2+} + H_2O_2 + H^+ \rightarrow Fe^{3+} + \bullet OH + H_2O \\ DNA + \bullet OH \rightarrow SSB \end{gathered} \quad (1)$$

A second H_2O_2 molecule can react at the same site to produce a DSB:

$$\begin{gathered} Fe^{2+} + H_2O_2 + H^+ \rightarrow Fe^{3+} + \bullet OH + H_2O \\ SSB + \bullet OH \rightarrow DSB \end{gathered} \quad (2)$$

Once DNA damage is recognized, signaling through ATM, ATR, and/or DNA-PK lead to increased activity of Chk1/2 and p53, resulting in cell cycle arrest and DNA damage repair [21-24]. If DNA damage cannot be repaired, cell death ensues via one of a number of possible mechanisms, including senescence, intrinsic apoptosis, mitotic catastrophy, autophagy, or necrosis [25-27].

2.1.3 Predictors of radiation therapy response

Early attempts at identifying predictive markers for radiation sensitivity in patient tumors began with the law of Bergonié and Tribondeau, who in 1906 stated that "X-rays act on cells inasmuch efficiently as cells have a greater reproductive activity, their karyokinetic fate is longer, their morphology and function are at least definitively fixed" [28, 29]. In 1927, Kimbrough et al. found that while 76% of cervical cancers of embryonal origin were sensitive to radiation therapy, only 51% of those with mature-cell composition were radiosensitive [30]. Kistner et al. later identified correlations between radiation sensitivity in cervical cancer with histologic grade [31]. These early studies validated the connection between tumor radiation response and cellular/molecular characteristics.

Later studies identified correlations between tumor radiation response and patient clinical factors. Kagan et al. found that success of radiation therapy for brain metastasis was 80% in patients with initial Karnofsky performance score greater than 60, but only 18% in patients with initial score of 50 or below [32]. Griffin et al. similarly found that initial Karnofsky score, as well as anatomical location and tumor staging, were independently prediction of radiation sensitivity in head and neck cancer [33]. Other

studies confirmed *in vitro* experiments dating back to 1936, when Mottram et al. showed that cancer cells in anaerobic conditions were more radiation resistant than in aerobic conditions [34]. Decreased patient survival was correlated with lower pretreatment oxygenation status by Nordsmark et al. in head and neck cancer patients, and later by Fyles et al. in cervical cancer patients [35, 36].

As the molecular basis of radiation therapy response and DNA damage repair in cancer cells was being studied, mutations in key enzymes which affect radiation response were identified. Homozygous mutations in the ATM gene (causing ataxia telangiectasia) and the LIG4 gene (causing DNA ligase IV syndrome) result in significant radiation sensitization, leading to an a surviving fraction at 2 Gy (SF2) of about 1-3%; however, the incidence of these mutations are extremely rare [37, 38]. Heterozygous mutations in BRCA1, BRCA2, and TP53 have also been associated with radiation sensitivity, although to a much lesser extent [37]. On the other hand, KRAS mutations in lung cancer and NFE2L2/KEAP1 mutations in head and neck cancers have been associated with increased resistance to radiation therapy; KRAS mutations result in increased osteopontin production which promotes stem-ness and represses DNA DSB induction, while activating mutations in NFE2L2 cause increased protection against oxidative damage from ionizing radiation [39, 40].

Currently, gene expression markers are most commonly studied for their diagnostic utility in predicting radiation response in tumors [8]. Choudhury et al. identified that pre-treatment expression of the MRE11 gene, involved in DNA damage repair, was predictive of survival following radiation therapy in bladder cancer patients [41]. Torres-Roca later developed a radiosensitivity panel composed of 10 genes (ABL1, AR, CDK1, HDAC1,

IRF1, JUN, PRRT2, RELA, STAT1, SUMO1) for prediction of progression-free survival in radiation-treatment patients compared to un-treated patients; this panel was validated in multiple cancer types, including breast cancer, NSCLC, and prostate cancer [42-45]. Most of these genes were previously implicated in the cell cycle, DNA damage response, immune system regulation, and tumorigenesis. Many other cancer-specific as well as pan-cancer gene expression signatures for the prediction of radiation response have since been developed, ranging in number of genes from tens to thousands. The RadiationGeneSigDB recently curated both oxic and hypoxic gene expression signatures for radiation response, comparing predictive performance and identifying similarities in gene expression biomarkers [46].

2.2 Redox Metabolism

2.2.1 Redox cofactors

2.2.1.1 NAD^+, NADH

Nicotinamide adenine dinucleotide (NAD) is an omnipresent molecule which acts as both an electron-carrying cofactor for oxidation–reduction reactions, as well as a substrate for many metabolic and signaling processes [47]. NAD metabolism has been implicated in several important biological processes, including energy production, cell signaling, and redox homeostasis. Many pathways in the NAD metabolic network are disrupted in cancer, including the production of NAD intermediates and consumption of NAD by signaling processes necessary for tumor cell survival [48-50].

Nicotinic acid mononucleotide (NAMN), a precursor of NAD^+, can be formed from either quinolinic acid (originating from tryptophan) in the de novo NAD^+ synthesis

pathway, or from nicotinic acid in the Preiss-Handler pathway (Figure 2-2) [51, 52]. NAMN is then converted to nicotinate adenine dinucleotide by nicotinamide mononucleotide adenylyl transferase (NMNAT), and then converted into NAD^+ by NAD^+ synthase. By utilizing salvage pathways, NAD^+ can also be produced using available precursors such as nicotinamide and nicotinamide riboside. The rate-limiting enzyme in these salvage pathways is nicotinamide phosphoribosyltransferase (NAMPT), which converts nicotinamide to nicotinamide mononucleotide [53].

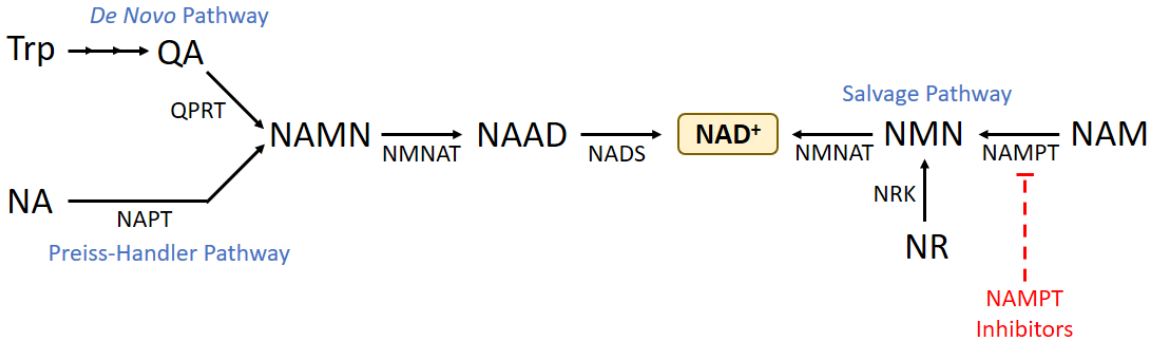


Figure 2-2. Many processes in the NAD^+ metabolic network are disrupted in cancer, including the production of NAD^+ intermediates and consumption of NAD^+ by signaling processes necessary for tumor cell survival.

Glucose oxidation supplies a significant amount of energy for the transfer of electrons from NAD^+ to NADH (Figure 2-3). Glycolysis via glyceraldehyde 3-phosphate dehydrogenase, oxidative decarboxylation of pyruvate via the pyruvate dehydrogenase complex, and the citric acid cycle via isocitrate dehydrogenase isoform 3 (IDH3), the α-ketoglutarate dehydrogenase complex, and malate dehydrogenase produce a combined 10 NADH per glucose molecule. In addition, beta oxidation of fatty acid molecules provides energy for NAD^+ reduction via 3-hydroxyacyl-CoA dehydrogenase. Other enzymes that are involved in the reduction of NAD^+ to NADH include glutamate dehydrogenase

(GLUD), malic enzyme (ME), and methylenetetrahydrofolate dehydrogenase 2-like protein (MTHFD2L). Whereas some enzymes are specific to the NAD^+ cofactor (including ME2), others can nonspecifically reduce both NAD^+ and $NADP^+$ (including GLUD1/2) with varying catalytic efficiencies. Cytosolic and mitochondrial NADH can be exchanged via the malate-aspartate and glycerol-3-phosphate shuttles [51]. The oxidized form of NAD^+ is maintained at much higher levels than reduced NADH (NAD^+/NADH ~ 3 to >100, depending on subcellular compartment as well as free vs. protein-bound states) [54, 55].

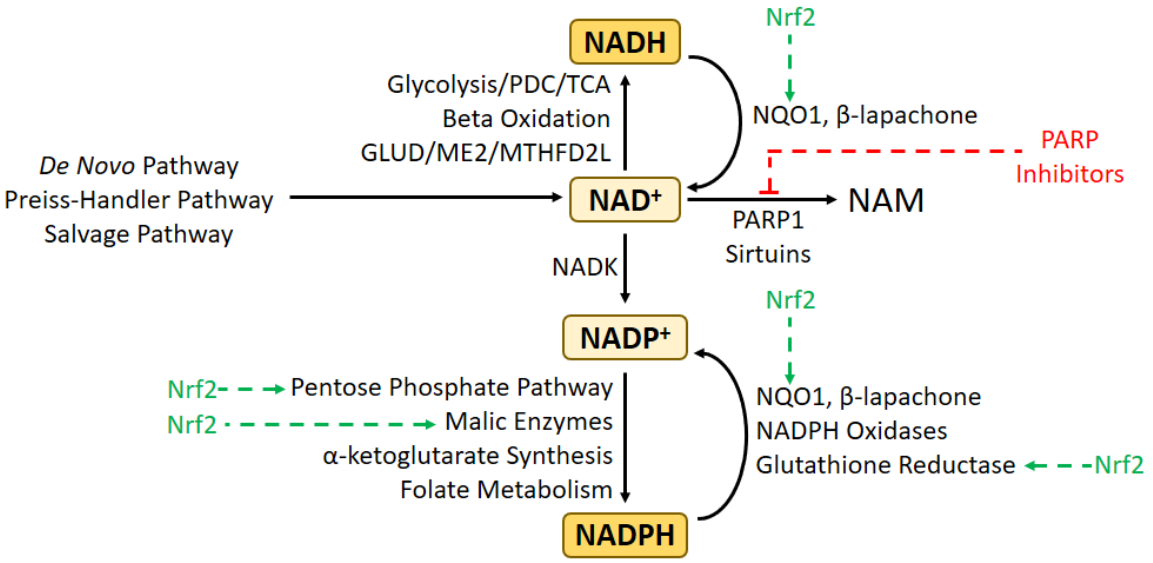


Figure 2-3. Major inputs and outputs to/from the cellular pools of NAD^+, NADH, $NADP^+$, and NADPH, including pathways and therapies pertinent to the cellular response to radiation therapy.

2.2.1.2 $NADP^+$, NADPH

NADPH is a ubiquitous metabolite implicated in over 900 reactions in human metabolism [7]. While NADH and NADPH are interchangeable as cofactors for a subset of these reactions, they mainly have distinct differences in function. The largely anabolic

couple [NADPH]/[$NADP^+$] (~100:1) is maintained out of equilibrium in cellular environments with [NADH]/[NAD^+] (1:10–1:1000), which is primarily used in catabolic processes [54]. NADPH/$NADP^+$ has a half-cell potential on the order of −400 mV and plays an essential role in limiting oxidative stress caused by reactive oxygen species (ROS); NADPH is a shared substrate for glutathione reductase and thioredoxin reductase, linking antioxidant defense and thiol-based signaling [56]. NADPH is also utilized as an electron donor in the synthesis of fatty acid, cholesterol, and nucleotide metabolites [57, 58].

$NADP^+$ is synthesized from NAD^+ through NAD kinase (Figure 2-3) [59]. Multiple metabolic reactions are implicated in the reduction of $NADP^+$ to NADPH (Figure 2-4). The pentose phosphate pathway contains two NADPH-producing reactions in series: G6PD, which is thought to carry the largest amount of NADPH-producing flux and is essential for cell survival, and PGD, which has also been implicated in tumor growth and redox metabolism [60, 61]. Differing isoforms of both malate enzyme and isocitrate dehydrogenase catalyze cytosolic (ME1, IDH1) and mitochondrial (ME2/3, IDH2) NADPH production utilizing substrates from the TCA cycle. Glutamine intake and conversion to glutamate allows for additional production of mitochondrial NADPH through glutamate dehydrogenase (GLUD1/2). Finally, multiple cytosolic (MTHFD1, ALDH1L1) and mitochondrial (MTHFD2, ALDH1L2) enzymes are involved in folate-dependent NADPH production. Because NADPH is involved in such a large number of reactions and pathways throughout the cellular metabolic network, it has not been traditionally considered in design of therapeutic strategies for selectively targeting radiation-resistant tumors.

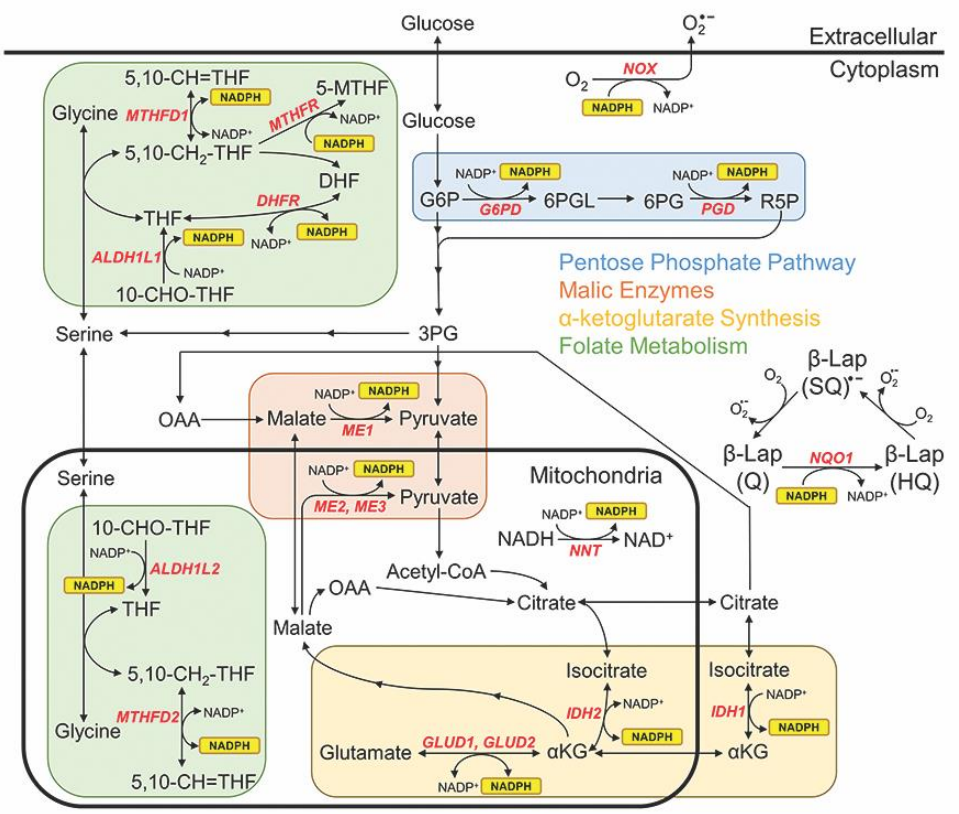


Figure 2-4. Major NADPH biogenesis pathways and reactions. Interconnection of well-studied biochemical reactions that convert oxidized NADP+ to reduced NADPH, labeled with their associated enzymes and grouped into biochemical pathways (*colored boxes*).

2.2.1.3 GSSG, GSH

Reduced glutathione (GSH) is a key antioxidant molecule, utilized by glutathione peroxidase (GPx) and glutaredoxin (Grx) systems for ROS scavenging. Oxidized glutathione (GSSG) is converted to GSH via glutathione reductase (GR) and NADPH. Normally the GSH/GSSG ratio exceeds 100:1, but decreases significantly with oxidative stress [62]. The glutathione half-cell reduction potential (E_{hc}) is often used as a

measurement of oxidative stress, and can be calculated from intracellular GSH and GSSG concentrations using the Nernst equation:

$$E_{hc} = E_{hc}^{o} - \frac{RT}{zF}\ln\left(\frac{[\mathrm{GSH}]^2}{[\mathrm{GSSG}]}\right) \quad (3)$$

where E_{hco} is the standard half-cell potential, R is the universal gas constant, T is the temperature, z is the number of electrons involved (2), and F is the Faraday constant [63]. Proliferating mammalian cells have glutathione E_{hc} values ranging from around -210 mV to -260 mV, whereas apoptotic cells are more oxidized with E_{hc} values around -170 mV [54].

De novo GSH synthesis is accomplished in a two-step processes utilizing enzymes glutamate-cysteine ligase (GCL) and GSH synthetase, consuming amino acids cysteine, glutamate, and glycine in the process [64]. Maintenance levels of intracellular cysteine are needed to prevent GSH efflux, extracellular degradation, and increased oxidative stress [65].

2.2.2 ROS-scavenging enzymes

Numerous ROS-scavenging systems are employed in different subcellular organelles to regulate ROS levels and oxidative stress. Superoxide dismutases (SOD) are enzymes which utilize metal ions to convert superoxide ($O_2^{\bullet-}$) to hydrogen peroxide (H_2O_2); different isoforms are present in the cytosol (Cu/ZnSOD SOD1), mitochondria (MnSOD SOD2), and extracellular space (Cu/ZnSOD SOD3) [66]. Catalase (CAT) is a major H_2O_2-scavening enzyme, especially in the peroxisome where catalase has increased localization and catalytic activity [67, 68]. Although catalase does not rely on NADPH or GSH as electron donors, it does require reduced NADPH to prevent oxidation of its active site by H_2O_2 [69]. Glutathione peroxidase (GPx) and glutaredoxin (Grx) both utilize GSH

to catalyze the reduction of free H_2O_2 and glutathionylated protein thiol groups caused by H_2O_2, respectively. Similarly, peroxiredoxin (Prx) and thioredoxin (Trx) both utilize the oxidized/reduced thioredoxin couple controlled by thioredoxin reductase (TR) and NADPH to reduce free H_2O_2 and oxidized protein thiol groups, respectively. Table 2-1 shows the alternative isoforms of these H_2O_2-scavening enzymes used in the cytosol and mitochondrion; having these different isoforms allows for differential regulation in response to compartment-specific oxidative stress [70].

Table 2-1. Alternative isoforms of cytosolic and mitochondrial GPx, Grx, Prx, and TR.

Enzyme	Cytosolic	Mitochondrial
GPx	GPX1, GPX2, GPX4	GPX1, GPX4
Grx	GLRX	GLRX2
Prx	PRDX1, PRDX2	PRDX3
TR	TXNRD1	TXNRD2, TXNRD3

2.2.3 Role of redox metabolism in tumor response to radiation therapy

Altered cellular function at any step in the recognition, response, or repair of DNA damage has been previously implicated in the sensitivity of cancer cells to ionizing radiation [37, 71, 72]. However, because the mechanism of action of radiation therapy involves ROS production and resultant DNA damage, redox biology and metabolism are believed to play an important role in tumor radiation sensitivity as well. Scavengers of ROS including superoxide dismutase (SOD), catalase (CAT), and glutathione peroxidase (GPx) has been previously shown to modulate intracellular ROS levels in cancer [73-75]. Positive correlation between GPx and radiation resistance was previously observed among human

cell lines; additionally, elevated expression of manganese superoxide dismutase (MnSOD) was observed in radiation-resistant mesothelioma cells [76]. MnSOD overexpression resulted in increased radiation resistance of stably transfected CHO cell lines [5]. Additionally, increased GPx and CAT activity was observed in radiation-resistant C3H/He Him mice compared to radiation-sensitive BALB/c/J Him mice. These findings suggest that ROS scavengers may be possible targets for improving radiation sensitivity.

The activation of both PARP1 and sirtuins by ionizing radiation causes a significant depletion of cellular stores of NAD^+ [77]. These stores can be replenished by increasing flux through the NAD^+ salvage pathways. However, consumption of ATP by NMNAT, one of the critical enzymes in NAD^+ salvage, causes significant depletion of ATP stores. In the presence of a low level of DNA damage, PARP activation may promote cell survival; on the other hand, the presence of widespread DNA damage causes PARP hyperactivation, severe NAD^+/ATP depletion, and programmed necrosis [78]. Radiation-resistant tumors are capable of maintaining adequate NAD^+ production while overcoming the damaging effects of radiation on DNA damage and ROS production [79]. Thus, it is expected that selectively targeting NAD^+ metabolism will sensitize tumor cells to ionizing radiation.

While NADPH metabolism is central to the ROS response and protection against oxidative stress, few enzymes involved in NADPH metabolism have been implicated in radiation response. IDH1 R132 mutations which cause oxidation of NADPH and production of the oncometabolite 2-hydroxyglutarate (2-HG) are common in low-grade gliomas and glioblastomas; these mutations have been associated with relatively favorable outcomes and improved sensitivity to radiation and chemotherapy, possibly due to NADPH depletion [80, 81]. G6PD overexpression was found to confer resistance to apoptosis

mediated by nitric oxide (NO), a reactive nitrogen species (RNS) which causes glutathione oxidation and NADPH consumption similar to ROS [82]. Additionally, inhibition of G6PD was found to sensitize cancer cells to treatment with cisplatin, a radiation-sensitizing chemotherapy [83]. Other studies have shown increased glutamine production in radiation-resistant cancers and that inhibiting glutamine metabolism improves radiation response, possibly implicating glutamate dehydrogenase as an important regulator of radiation response [84, 85].

2.3 Radiation-Sensitizing Chemotherapies

2.3.1 General principles

The use of chemotherapy in conjunction with radiation treatment to enhance radiation sensitivity began in 1958, when Heidelberger et al. showed increased response in preclinical studies of solid tumors when combining radiation treatment with 5-fluorouracil [86]. In 1974, Nigro and colleagues led a pioneering study where three out of three patients with anal cancer who were treated with a combination of 5-fluorouracil, mitomycin D, and radiation therapy achieved complete responses [87]. Today, chemotherapy-radiation combination therapies are used for a number of cancer types, including bladder, cervical, esophageal, gastric, glioblastoma, head and neck, NSCLC, and rectal cancers [88-95].

The mechanism of action of many radiation-sensitizing chemotherapies involves the induction of DNA damage or inhibition of DNA damage repair [96]. Examples of DNA-targeting radiosensitizers include cisplatin, an alkylating agent which inhibits DNA synthesis by forming DNA cross-links, cyclophosphamide, another alkylating agent, and doxorubicin, an inhibitor of topoisomerase II and DNA intercalating agent [97-99]. Other

types of radiation-sensitizing therapies include those which inhibit pro-survival pathways (e.g. EGFR inhibitors), target tumor vasculature (e.g. VEGF inhibitors), or alter cell cycle checkpoints (e.g. AZD7762) [100-102].

Chemotherapies which modulate redox biology and metabolism are also garnering attention for their utility as radiosensitizers [103]. Rischin et al. found that combining cisplatin and radiation with tirapazamine, a compound which is metabolized and causes ROS formation in anaerobic conditions, resulting in borderline significant failure-free survival rates among patients with advanced head and neck cancer [104]. Shaw et al. found that combining radiation therapy with RSR13, a hemoglobin-binding inhibitor which enhances oxygen delivery to tissues and thereby increases radiation response, resulted in significant benefit for patients with brain metastases from breast cancer [105]. While novel redox-modulating radiosensitizers are showing promise, many conventionally-used chemotherapies have effects on redox metabolism which cause oxidative stress and may subsequently sensitize to radiation treatment; however, these alternative redox-based mechanisms of radiosensitization are less characterized than their commonly-attributed mechanisms of action involving DNA damage [6].

2.3.2 Redox metabolism of radiation-sensitizing chemotherapies

2.3.2.1 Cisplatin

Cisplatin's primary redox-based mechanism of action is its binding to mitochondrial DNA (mtDNA), causing impaired synthesis of electron transport chain proteins, ETC malfunction, and subsequent increases in superoxide production (Figure 2-5) [106-108]. Cisplatin is imported into cells via the copper transporter CTR1 [109]. ATP-catalyzed efflux of the drug can occur via the breast cancer resistance protein (BRCP) or

copper efflux transporters ATP7A and ATP7B [110-112]. After import, cisplatin is transported to the mitochondria via the copper chaperone Cox17 [113]. Cytosolic cisplatin can be cleared by both glutathionylation as well as interaction with metallothioneins, a family of trace metal and free radical scavenging proteins [114, 115]. GSH consumption from cisplatin glutathionylation may also contribute to redox-based sensitization of cancer cells to ROS-generating ionizing radiation.

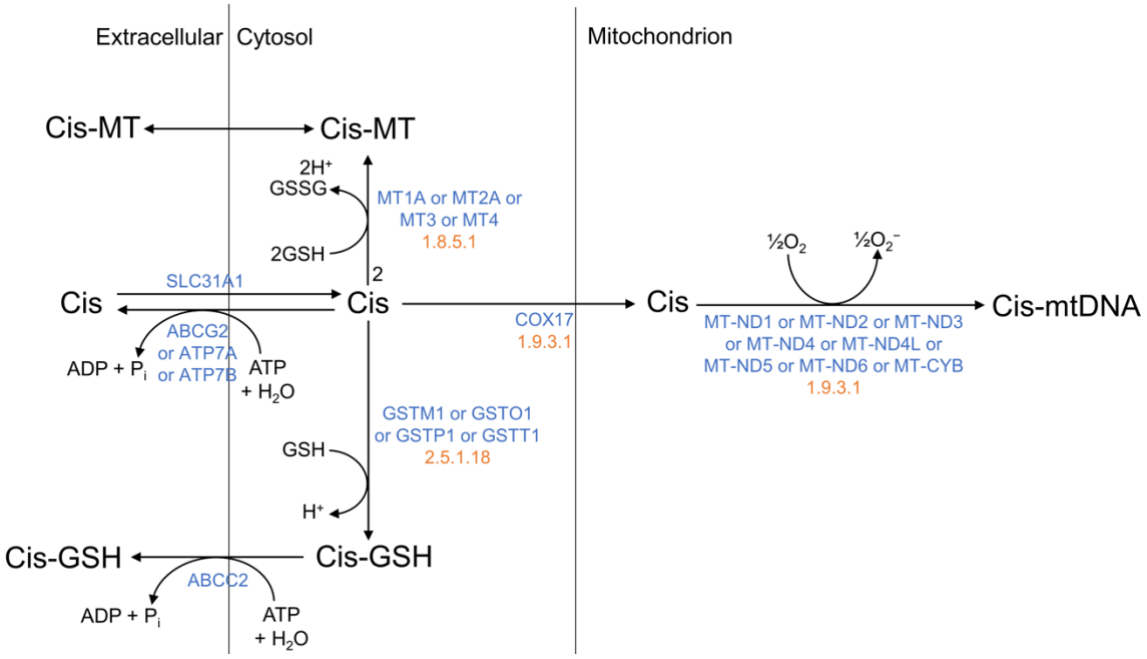


Figure 2-5. Overview of cisplatin metabolism, including import/export reactions, mechanism of action, and drug clearance.

2.3.2.2 Cyclophosphamide

While the primary mechanism of cyclophosphamide involves production of phosphoramide mustard and subsequent DNA cross-linking, cyclophosphamide metabolism involves the utilization of multiple redox cofactors, including NADH, NADPH, and GSH; the consumption of these cofactors may contribute to

cyclophosphamide's radiation-sensitizing effects (Figure 2-6) [116-118]. Following import, cyclophosphamide efflux can occur via multidrug resistance protein 1 (MDR1) or breast cancer resistance protein (BRCP) [119, 120]. Cyclophosphamide can be metabolized into 4-hydroxycyclophosphamide (4OH-CPA) or into dechloroethylcyclophosphamide (DCE-CPA) and chloroacetaldehyde (CAA) by NADPH-catalyzed cytochrome P450 activity. 4OH-CPA and related metabolite aldophosphamide (ALDPh) are metabolized using NAD_+ cofactors by alcohol dehydrogenase and aldehyde dehydrogenase, respectively [121]. NADPH is also used in ALDPh metabolism by aldo-keto reductase. Finally, end products phosphoramide mustard and acrolein are glutathionylated and exported [122].

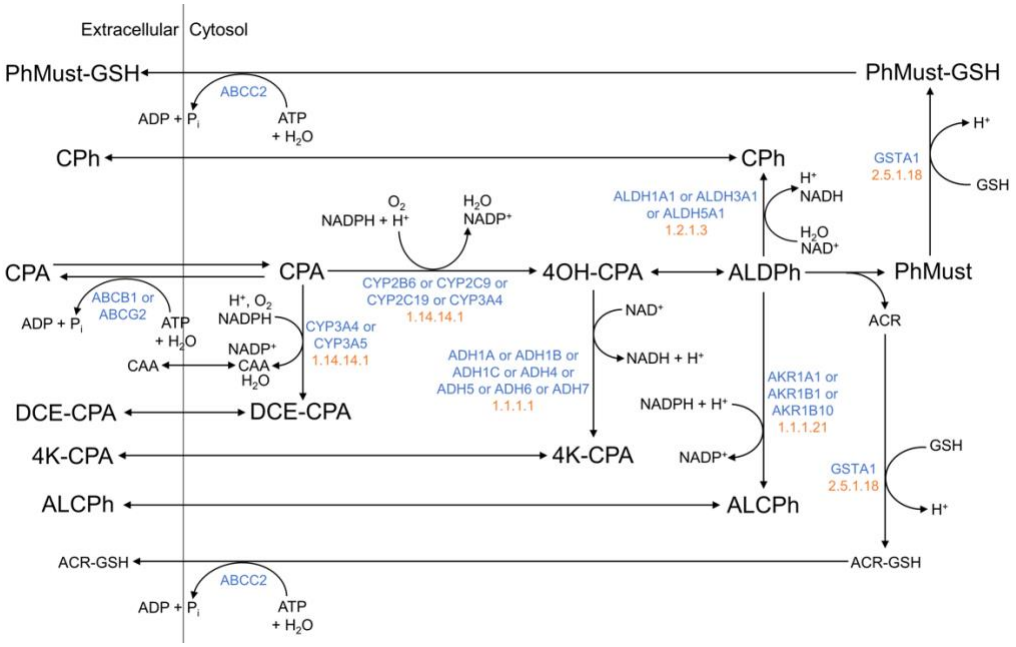


Figure 2-6. Overview of cyclophosphamide metabolism, including import/export reactions, mechanism of action, and drug clearance.

2.3.2.3 Doxorubicin

Doxorubicin undergoes a one-electron redox cycling mechanism between quinone and semiquinone forms, resulting in both superoxide production and NADPH consumption which may contribute to radiation sensitization (Figure 2-7) [123-125]. Import of doxorubicin occurs via the organic cation transporter SLC22A16 [126]. Doxorubicin efflux can occur by either RLIP76 or ATP-dependent ABC-family transporters [127]. The quinone form of doxorubicin can be converted to its semiquinone form through either nitric oxidase synthase 3 (NOS3, eNOS) or cytochrome P450 oxidoreductase (POR) [128-130]. These reactions differ in the stoichiometry of NADPH consumption as well as the production of the RNS nitric oxide (NO); the relative importance of these two reactions to doxorubicin metabolism and radiation sensitization has not been determined. Following reduction, the semiquinone form spontaneously converts back to the quinone form, producing superoxide in the process. Doxorubicin quinone is cleared by carbonyl reductase, aldo-keto reductase, and cytochrome P450 reductase, all of which consume NADPH [131].

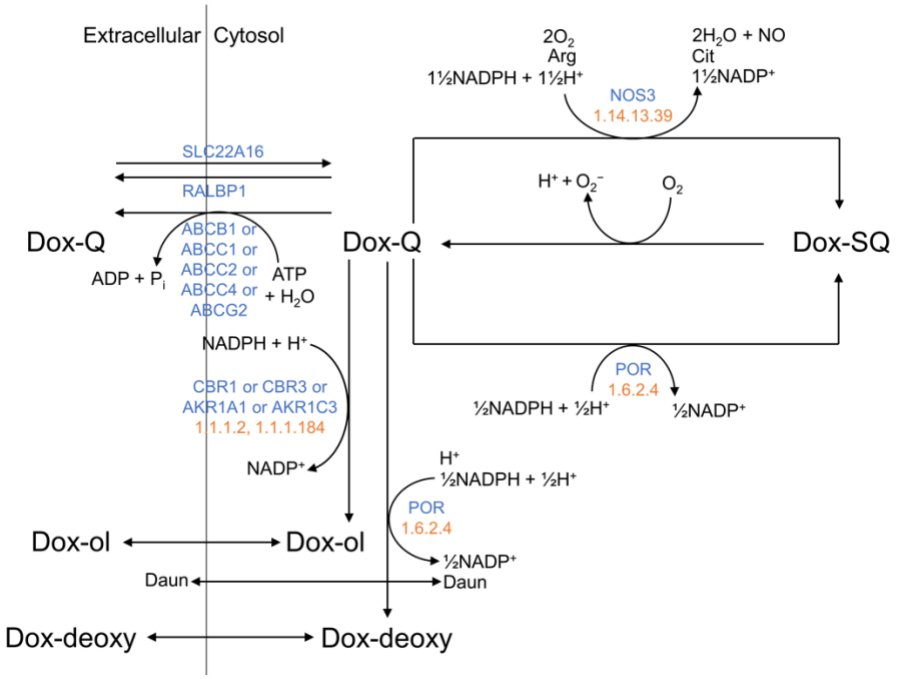


Figure 2-7. Overview of doxorubicin metabolism, including import/export reactions, mechanism of action, and drug clearance.

2.4 Flux Balance Analysis

Flux Balance Analysis (FBA) is a constraint-based computational modeling approach for predicting steady-state metabolic fluxes at a genome-scale [9]. Reaction fluxes are predicted by combining (1) a genome-scale metabolic network reconstruction for the organism of interest, including metabolites, reactions, and the genes associated with metabolic enzymes; (2) constraints imposed on minimum and maximum reaction fluxes, which are often implemented using biological information on the sample of interest; and (3) an objective function, which is a reaction flux or function of multiple fluxes to be maximized within the metabolic network.

2.4.1 Mathematical representation

2.4.1.1 Flux Balance Analysis (FBA)

A metabolic network can be represented by a stoichiometric matrix *S* of size *m* x *r*, where *m* and *r* are the number of metabolites and reactions in the network, respectively. Entry S_{ij} is equal to the stoichiometric coefficient of metabolite *i* in reaction *j* ($S_{ij} < 0$ for reactants, > 0 for products, and = 0 if metabolite *i* is not involved in reaction *j*). The relationship between reaction fluxes and metabolite concentrations in the network can be written as:

$$S\mathbf{v} = \frac{d\mathbf{m}}{dt} \tag{4}$$

where **v** is a *r* x 1 vector of reaction fluxes, and **m** is a *m* x 1 vector of metabolite concentrations. In flux balance analysis, the steady state is assumed (metabolite concentrations do not change with time), changing Equation 4 to:

$$S\mathbf{v} = \mathbf{0} \tag{5}$$

Each reaction flux v_j is also constrained by lower and upper bounds:

$$v_{min,j} \leq v_j \leq v_{max,j} \tag{6}$$

The solutions to Equations 5 and 6 that maximize a particular objective function $f(\mathbf{v})$ are chosen. Thus, the FBA problem can be represented as a Linear Programming (LP) problem:

$$\begin{gathered} \max f(\mathbf{v}) \\ \text{subject to } S\mathbf{v} = \mathbf{0} \\ \text{and } v_{min,j} \leq v_j \leq v_{max,j} \end{gathered} \tag{7}$$

Solving this LP problem provides us with the maximized value f_{max} of the objective function. LP problems can be solved computationally using commercially-available optimization software including Gurobi.

2.4.1.2 Parsimonious Flux Balance Analysis (pFBA)

Despite maximizing an objective function, Equation 7 can still result in more than one possible solution for the flux vector **v** which satisfies network stoichiometry and model constraints. One approach to obtain a single representative flux vector **v** is parsimonious FBA (pFBA), which also minimizes the total sum of fluxes throughout the metabolic network:

$$\min \sum_{j=1}^{r} |v_j| \text{ s.t. } f(\mathbf{v}) = f_{max} \tag{8}$$

where *r* is the number of reactions in the metabolic network [132]. The biological reasoning behind pFBA is that metabolic networks are optimized to maximize their objective while utilizing the minimum amount of metabolic resources (including substrates and metabolic enzymes) necessary. While having a single representative flux vector facilitates downstream analysis of model outputs, this single vector may not be representative of the entire solution space characterized by Equation 7.

2.4.1.3 Flux Variance Analysis (FVA)

Flux variance analysis (FVA) allows for the calculation of the minimum and maximum allowable fluxes through each metabolic reaction v_j that still maintain the maximum possible objective value as calculated in Equation 7:

$$\min/\max v_j \text{ s.t. } f(\mathbf{v}) = f_{max} \tag{9}$$

This provides an overview of the range of possible flux values that each reaction can obtain, which is more representative of the entire solution space than the single flux vector obtained from pFBA. While the entire distribution of values for each reaction flux v_j is not obtained, the average flux through each reaction can be estimated by taking the mean of

the minimum and maximum values from Equation 9. To perform FVA, two LP problems must be solved per metabolic reactions; for networks with tens-of-thousands of reactions, performing FVA across the entire metabolic network takes significantly more computational time than FBA.

2.4.1.4 Random Sampling

An optimal representation of the possible solution space from Equation 7 is obtained from random sampling, where a user-specified number of points are samples within the solution space to obtain a distribution of flux values for each metabolic reaction. Software such as optGpSampler has been developed for uniform sampling of the solution space for constraint-based metabolic models [133]. However, the computation time of random sampling increases exponentially with the size of the metabolic network; for current metabolic reconstructions with tens-of-thousands of reactions, performing random sampling for hundreds of samples and thousands of perturbations per sample (for example, with simulated gene knockout screens) is computationally infeasible.

2.4.2 Genome-scale metabolic reconstructions

Genome-scale metabolic reconstructions are comprehensive representations of all known metabolic reactions, metabolites, and enzyme-encoding genes within the entire metabolic network of a particular organism [134, 135]. These reconstructions are developed through iterative processes involving the identification of candidate metabolic features, refinement based on available biochemical information, conversion into a format which can be used in computational models, and evaluation of resulting metabolic predictions to ensure accuracy and compatibility with other metabolic features [136]. The latest human metabolic reconstruction, Recon3D, incorporates 8,401 metabolites, 13,547

reactions, and 3,268 genes [7]. Metabolites and reactions are represented mathematically in the stoichiometric matrix, as explained above. Genes are represented in gene-protein-reaction rules (GPR) which, for each reaction, shows which genes encode the metabolic enzymes which catalyze the associated metabolic reaction. Gene names are given as either the NCBI gene symbol or gene ID number. GPR's are written as Boolean functions. For example, if the enzymes encoded by either Gene1 or Gene2 can catalyze the reaction on their own, the GPR would be "Gene1 OR Gene2". If both enzymes are needed to catalyze the reaction, the GPR would instead be "Gene1 AND Gene2". When incorporating expression data into FBA models, "OR" statements are handled by adding the expression values from both genes, and "AND" statements are handled by taking the minimum expression value among both genes. Finally, each reaction has an associated enzyme commission (EC) number, which provides information about the catalytic activity of the particular enzyme [137].

2.4.3 FBA objective functions

An objective function is maximized in Equation 7 to obtain a single (or significantly smaller subset) of feasible flux vectors which satisfy network stoichiometry and reaction constraints. This objective function is often used to represent a biological "objective" that the organism/cell/sample under study is attempting to maximize physiologically [138]. To this end, the most commonly used objective function has been maximizing the production of biomass, which represents all of the carbohydrate, protein, lipid, nucleotide, and other metabolites which a cell or organism needs to survive and grow [139]. The use of the biomass objective function as a proxy for optimizing cellular growth has been applied for FBA models of both microorganisms as well as cancer cells which demonstrate sustained

proliferation; however, its use for most normal human tissues which show minimal proliferation is likely unwarranted [140, 141]. Other physiological objective functions which have been used previously include the production of metabolic cofactors, including ATP, NADH, and NADPH, which are used ubiquitously throughout the metabolic network for anabolic and redox reactions [142-144].

Alternative approaches for objective functions have included the incorporation of experimental data from the sample of interest into the objective function itself. Lee et al. used an objective function which maximized the correlation between experimentally-measured gene expression data from *Saccharomyces cerevisiae* and predicted FBA fluxes to improve the accuracy of metabolic flux predictions [145]. Montezano et al. developed an objective function as the linear combination of reaction fluxes with weights proportional to measured proteomics data from *Mycobacterium tuberculosis* exposed to the antibiotic mefloquine [146]. It is yet unclear whether the incorporation of experimentally-measured gene/protein expression data as constraints on metabolic fluxes, as modifiers of the objective function, or some combination of the two, results in the most accurate metabolic predictions.

2.4.4 Implementation of model constraints using biological data

FBA models that are personalized to individual cell types or samples of interest are most commonly developed by incorporating genome-scale experimental data into the minimum and maximum flux constraints for each reaction (Equation 6). Transcriptomic data, either from microarrays or RNA-seq, is most often used due to its much wider availability for samples of interest compared to genome-scale proteomic or metabolomic data [11]. Numerous algorithms for the implementation of transcriptomic data into reaction

constraints have been developed, all with associated advantages and disadvantages compared to other methods. One of earliest and simplest methods was developed by Akesson et al., who set reaction fluxes to zero if the associated gene expression was not detected from the utilized microarray platform [147]. Similarly, the GIMME algorithm developed by Becker et al. implemented a user-defined gene expression threshold to determine whether individual metabolic reactions were to be set as active (reaction allowed to carry flux) or inactive (flux set to zero) [148]. iMAT, developed by Shlomi et al., utilizes a slightly different approach by discretizing gene expression into high, medium, and low expression; reaction fluxes associated with high gene expression are maximized, whereas fluxes associated with low gene expression are minimized [149]. All of the aforementioned algorithms rely on arbitrary threshold of gene expression data which, although simple to understand and implement, does not accurately reflect biological regulation of metabolic reactions.

E-Flux, developed by Colijn et al., constraints the upper flux bound of each reaction relative to the expression of the associated gene instead of implementing binary on/off states; however, a threshold is still utilized to determine whether loose or tight constraints are used for reactions with gene expression higher or lower than the threshold, respectively [150]. PROM, developed by Chandrasekaran et al., constrains reaction flux bounds by calculating the likelihood that the gene is active relative to the expression of the gene's transcription factor, and sets upper flux bounds in proportion to this calculated probability [151]. To avoid setting arbitrary thresholds, MADE, developed by Jensen et al., compares multiple gene expression datasets and assesses statistically significant differences to set on/off states of reactions for different conditions [152]. These newer methods provide more

sophisticated approaches towards integrating gene expression data into reaction upper bounds; however, they still rely on either thresholding of expression values or comparison to other datasets to set model constraints.

Other biological data types other than gene expression have been explored for utilization to set metabolic reaction constraints. INIT, developed by Agren et al., uses qualitative confidence scores derived from proteomic data from the Human Protein Atlas to maximize flux through high-expression reactions and minimize flux through reactions with absent expression [153]. Approaches such as those from Hoppe et al. and Yizhak et al. implement metabolomic data from *E. coli* to apply additional constraints on plausible metabolic concentration ranges or kinetically-derived flux estimates [154, 155]. Finally, dynamic flux balance analysis (dFBA), developed by Mahadevan et al., utilizes time-resolved biochemical parameters to predict changes in metabolic fluxes and concentrations over time [156]. Because of the constraint-based nature of FBA models, additional constraints based on multiple biological data types will further constrain flux solution spaces and will ultimately yield more accurate predictions of metabolic phenotypes in samples of interest.

CHAPTER 3. FBA MODELS OF NAD(P)H-DRIVEN ß-LAPACHONE SENSITIZATION

3.1 Introduction

A feature of NAD(P)H utilization is the phenomenon of redox cycling, in which two coupled reduction and oxidation reactions can perpetuate a sustained directional cyclic flux of a quinone substrate between quinone and semiquinone (reduced) forms due to an ample source of electrons. Noted as early as 1968 in *Escherichia coli* metabolism of streptonigrin [157], NADPH-dependent redox cycling has been attributed to acetaminophen hepatotoxicity and cardiac lethality of anthracyclines *via* cytochrome P450 reductase [158]. Systems biology modeling of this latter mechanism identified the role of NADPH availability in modulating the doxorubicin concentration-dependent switch in ROS formation [125]. Model-predicted control *via* glucose-6-phosphate dehydrogenase (G6PD) was experimentally tested through inhibition of the enzyme and confirmed to alter cell line-specific changes in drug sensitivity.

A distinct but related mechanism of drug redox cycling is observed through NAD(P)H: quinine oxidoreductase 1 (NQO1) bioactivatable compounds. Drugs such as β-lapachone and deoxynyboquinone rely upon two-electron transfer *via* NQO1 to catalyze interconversion between quinone, hydroquinone, and semiquinone forms, expending one NAD(P)H molecule per quinone oxidation to generate two molecules of superoxide [159]; they are the only known quinones to cycle through this mechanism. Tumor cells, which

* As published in (1) Lewis et al., Antioxid Redox Signal, 2018; (2) Lewis et al., Semin Radiat Oncol, 2019

typically express higher ratios of NQO1 to catalase (CAT) than nonmalignant tissue, are known to redox cycle in a futile manner to generate 120 moles of superoxide in 2 min for every mole of β-lapachone [160]. Cellular superoxide dismutase (SOD) enzymes convert superoxide to the relatively more stable hydrogen peroxide (H_2O_2). Elevated H_2O_2 (>300 μ*M*) leads to extensive base damage, single-strand DNA lesions, and eventually double-strand lesions of DNA. Furthermore, the high rate of NAD(P)H consumption contributes to depletion of the nicotinamide pool by subsequent poly(ADP-ribose)polymerase-1 (PARP1) hyperactivation, with cell death proceeding through caspase-independent programmed necrosis [161-163].

NADPH is also oxidized to $NADP^+$ by ribonucleotide reductase to maintain dNTP pools for DNA repair and cell survival following ionizing radiation [164]. A synergistic effect is reported when combining sublethal doses of β-lapachone with radiation therapy in $NQO1^+$ human prostate and head and neck cancers [165-167], as well as in non-small cell lung cancers. Increased consumption of NAD^+ and ATP, enhanced DNA double-strand breaks, and prolongation of γ-H2AX were attributed, in the combined therapy, to result in head and neck squamous cell carcinoma (HNSCC) tumor regression. The competition for NADPH as a substrate for futile redox cycling, reduction of antioxidant defense systems (thioredoxin and glutaredoxin), PARP activity, and as a cofactor for dNTP synthesis upon radiation and chemotherapeutic insult in high NQO1/CAT tumor cells results in selective cytotoxicity.

Recent bioinformatic analyses have indicated that genetic or pharmacological suppression of intracellular [NADPH] through the NAMPT or glutamine biosynthesis pathway can confer enhanced sensitivity to β-lapachone by reduced antioxidant defenses

and increased ROS, thereby lowering the *in vivo* dose needed for selective tumor cytotoxicity [168-170]. This result suggests that strategically targeting NADPH-biosynthesis pathways concomitant with NQO1-bioactivated futile redox cycling for therapeutic design may be advantageous. Flux balance analysis (FBA) is an optimal tool to systematically analyze the genes which contribute to whole-cell NAD(P)H supply and demand across the entire metabolome, and how this may impact the response to redox-cycling therapies. To evaluate the role of global NAD(P)H production on ß-lapachone sensitization, as well as to improve upon existing FBA models, we have developed a human genome-scale metabolic model that incorporates quantitative transcriptomic, kinetic, thermodynamic, and metabolite concentration constraints. These constraints have the potential to greatly improve the accuracy and cell-type specificity of predicted flux distributions associated with redox metabolism. The goal of our model development was to compare intrinsic metabolic changes in matched head and neck cancer cell lines that may yield phenotypic differences in sensitivity to NAD(P)H: quinone oxidoreductase 1 (NQO1)-dependent drugs such as the chemotherapeutic compound β-lapachone that relies on redox cycling, ROS production, and nicotinamide depletion for its cytotoxic effects.

3.2 Methods

3.2.1 Computational methods

3.2.1.1 Redox-based changes to Recon 2 network

Recon 2.04 was used as the core metabolic network to build our SCC-61 and rSCC-61 models [134, 135]. To address missing and inaccurate redox-based reaction information within the Recon 2 metabolic network, we have made the following changes (for the

reactions below, c = cytoplasm, e = extracellular, n = nucleus, and r = endoplasmic reticulum):

1. We split reaction FTHFDH (NADPH-producing reaction originally catalyzed by ALDH1L1 and ALDH1L2) into two separate cytosolic and mitochondrial reactions. The cytosolic reaction is catalyzed by ALDH1L1, and the mitochondrial reaction is catalyzed by ALDH1L2.

   Originally:

   a. 10-CHO-THF[c] → THF[c], GPR = ALDH1L1 or ALDH1L2, EC = 1.5.1.6

   Updated:

   a. 10-CHO-THF[c] → THF[c], GPR = ALDH1L1, EC = 1.5.1.6
   b. 10-CHO-THF[m] → THF[m], GPR = ALDH1L2, EC = 1.5.1.6

2. We removed gene MTHFR from reaction MTHFD (NADPH-producing reaction originally catalyzed by MTHFR and MTHFD1). MTHFR catalyzes the reaction converting 5,10-CH_2-THF to 5-MTHF, which is already included in the Recon 2 network.

   Originally:

   a. 5,10-CH_2-THF[c] → 5,10-CH=THF[c], GPR = MTHFR or MTHFD1, EC = 1.5.1.5

   Updated:

   a. 5,10-CH_2-THF[c] → 5,10-CH=THF[c], GPR = MTHFD1, EC = 1.5.1.5

3. We added reactions catalyzed by isoforms of NADPH oxidase.

   Originally: None

Updated:

a. NADPH[c] + 2 O_2[c] → 2 $O_2^{\bullet-}$[e] + $NADP^+$[c] + H^+[c], GPR = NOX1 or CYBB or NOX3 or NOX5, EC = 1.6.3.1
b. NADPH[n] + 2 O_2[n] → 2 $O_2^{\bullet-}$[c] + $NADP^+$[n] + H^+[n], GPR = NOX4, EC = 1.6.3.1
c. NADPH[r] + 2 O_2[r] → 2 $O_2^{\bullet-}$[c] + $NADP^+$[r] + H^+[r], GPR = NOX4, EC = 1.6.3.1

Note: The superoxide dismutase-catalyzed reaction is already in Recon 2, allowing for the production of hydrogen peroxide from superoxide.

3.2.1.2 Objective function

Because FBA models produce a steady-state solution, there cannot be net consumption or production of any metabolites in the model. However, adding an artificial demand reaction allows for net metabolite consumption or production through existing model reactions. In order to maximize the flux through existing Recon 2 reactions that convert $NADP^+$ into NADPH within the cytoplasm and mitochondria, a new reaction is added with the following stoichiometry:

$$\mathrm{NADPH[c] + NADPH[m] \rightarrow NADP^+[c] + NADP^+[m] + H^+[c] + H^+[m]} \quad (10)$$

where c = cytoplasm and m = mitochondria. Maximizing this reaction as the objective function in turn maximizes the net conversion of $NADP^+ + H^+$ into NADPH through all other existing model reactions. This allows for the calculation of net $NADP^+ \rightarrow$ NADPH flux (equal to the flux through the maximized objective function), as well as the flux through all Recon 2 reactions that convert $NADP^+$ into NADPH.

Similarly, the following objective function is used for analyzing the net conversion of NAD_+ to NADH within the cytoplasm and mitochondria:

$$\mathrm{NADH[c] + NADH[m] \rightarrow NAD^+[c] + NAD^+[m] + H^+[c] + H^+[m]} \quad (11)$$

However, the sole maximization of $NADP_+ \rightarrow$ NADPH flux (or $NAD_+ \rightarrow$ NADH flux) may result in biologically non-viable solutions since it does not take into consideration the production of energy sources such as ATP that are necessary for cell survival. To ensure that biologically viable outcomes are obtained, we first maximize cellular ATP production through the following objective function:

$$\begin{aligned}\mathrm{ATP[c] + ATP[m] + H_2O[c] + H_2O[m]} \\ \rightarrow \mathrm{ADP[c] + ADP[m] + P_i[c] + P_i[m] + H^+[c] + H^+[m]}\end{aligned} \quad (12)$$

where c = cytoplasm and m = mitochondria. After calculating the maximum cellular ATP production, we set the lower bound of this reaction to 50% of the maximum value. This creates an ATP demand where only solutions that result in at least 50% of maximal ATP production are considered. While imposing this ATP demand, we then maximize cellular conversion of $NADP_+$ to NADPH (NAD_+ to NADH) or through the objective function given above.

3.2.1.3 Thermodynamic constraints

In order to include thermodynamic constraints within the FBA model, all reversible reactions (reactions that can carry a net flux in either the forward or reverse direction) within the model had to be broken down into two separate reactions - one which only runs in the original forward direction, and one which only runs in the original reverse direction. All reactions within the model now only have a forward direction. This allows us to set all $v_{min,j}$ values to zero.

Two additional sets of equations were added to include thermodynamic constraints within the FBA model. The first is of the form:

$$v_j \leq b_j v_{max,j} \tag{13}$$

where b_j is a Boolean variable representing the thermodynamic status of reaction j. If reaction j is thermodynamically feasible, then $b_j = 1$, and $v_j \leq v_{max,j}$ (Equation 13 with $v_{min,j} = 0$). If reaction j cannot carry a net flux due to thermodynamic constraints, then $b_j = 0$, and $v_j \leq 0$ (since $v_{min,j} = 0$, this means that v_j must equal zero).

To determine the value of b_j, another set of equations is included of the form:

$$Kb_j \leq K - \Delta G_j \tag{14}$$

where ΔG_j is the change in Gibbs free energy of reaction j, and K is an arbitrarily large constant that must be set greater than the largest $|\Delta G_j|$ value in the model. A chemical reaction j can carry a net flux if $\Delta G_j < 0$. If $\Delta G_j < 0$, then K - ΔG_j > K, and Equation 14 will hold whether $b_j = 0$ or $b_j = 1$. Any thermodynamically-feasible reaction that may contribute to the maximization of our objective function will choose to have $b_j = 1$ if possible, in order to carry a nonzero flux. However, if $\Delta G_j > 0$, K - ΔG_j < K, and Equation 14 will only hold if $b_j = 0$. Therefore, this reaction cannot carry a nonzero flux.

The ΔG_j of non-transport reactions can be calculated as:

$$\Delta G_j = \Delta G_j^o + RT \sum_{i=1}^{m} S_{ij} \ln x_i \tag{15}$$

where ΔG_{jO} is the standard change in Gibbs free energy of the reaction, R is the gas constant, T is the temperature, S_{ij} is the stoichiometric coefficient of metabolite i in reaction j, and x_i is the concentration of metabolite i. For reactions involving transport of metabolites between cellular compartments, ΔG_j is:

$$\Delta G_j = \Delta G_j^o + RT \sum_{i=1}^{m} S_{ij} \ln x_i + \sum_{i=1}^{m} c_i S_{\text{trans},ij} FV \quad (16)$$

where c_i is the charge of metabolite i, $S_{trans,ij}$ is the transport stoichiometric coefficient of metabolite i in reaction j (representing how many molecules of the metabolite is being transported between compartments), F is the Faraday constant, and V is the potential of the membrane separating the transport metabolites. V_m is estimated using the Goldman-Hodgkin-Katz voltage equation:

$$V = \frac{RT}{F} \ln\left(\frac{P_{\text{Na}^+}[\text{Na}^+]_{\text{out}} + P_{\text{K}^+}[\text{K}^+]_{\text{out}} + P_{\text{Cl}^-}[\text{Cl}^-]_{\text{in}}}{P_{\text{Na}^+}[\text{Na}^+]_{\text{in}} + P_{\text{K}^+}[\text{K}^+]_{\text{in}} + P_{\text{Cl}^-}[\text{Cl}^-]_{\text{out}}}\right) \quad (17)$$

where P_{Na+}, P_{K+}, and P_{Cl-} are the relative permeabilities of Na^+, K^+, and Cl^- across the membrane (assigned to 1, 5, and 80, respectively).

Estimated ΔG_{jO} values for many reactions in the Recon 2 model were available from the Virtual Metabolic Human database [135]. For any reaction where ΔG_{jO} was not available, the value ΔG_{jO} = -999999 was used so that reactions with no available thermodynamic data were not made thermodynamically infeasible.

Gurobi 6.5 optimization software was used to solve the Mixed-Integer Linear Programming problem represented by Equations 7, 13, and 14.

3.2.1.4 Proteomic and kinetic constraints

For reactions in the Recon 2 model with an associated enzyme commission [134] number and gene-reaction rule, the upper bound on the reaction flux was set to be:

$$v_{\text{max},j} = k_{\text{cat},j}[\text{E}]_j \quad (18)$$

where k_{cat} is the turnover number (1/hr) and $[E]_j$ is the total abundance (mmol/gDW) of the enzyme catalyzing reaction j. This is equivalent to the formula for v_{max} in Michaelis-Menten kinetics. Reactions that are missing either an EC number or gene-reaction rule

(either because the reaction is not enzyme-catalyzed or due to missing information) were unconstrained in an upper bound by assigning an artificially high $v_{max,j}$ = 999999.

3.2.1.5 Proteomic and transcriptomic data

To estimate protein expression values for SCC-61 and rSCC-61 cell lines, gene expression data for these cells [171] was compared to gene expression data for CAL-27 cells measured using the same gene expression microarray [172]. Data from three samples each of SCC-61, rSCC-61, and CAL-27 were log2 transformed and normalized using z-score normalization. ComBat was then used to remove batch effects between the SCC-61/rSCC-61 data and CAL-27 data [173]. For each Illumina probe, the mean expression value for each cell line was calculated. The fold change comparing SCC-61 to CAL-27, as well as comparing rSCC-61 to CAL-27, was calculated for each probe. The fold changes for all Illumina probes for a particular gene were averaged. Protein expression values in ppm units were calculated for SCC-61 and rSCC-61 by multiplying CAL-27 expression values from the ProteomicsDB database by the SCC-61/CAL-27 and rSCC-61/CAL-27 fold changes, respectively [174]. The rSCC-61 and SCC-61 protein concentration values estimated in this manner were then compared with stable isotope labeling of amino acids in cell culture (SILAC) proteomic data acquired by mass spectrometry [175]. The experimentally obtained values provide independent confirmation of the relative rSCC-61/SCC-61 values on 8.7% coverage of the proteome needed for the FBA models.

Missing protein expression values were imputed by performing k-nearest neighbors between data from all available tissues and cell lines from the GeneCards database, as well as CAL-27, SCC-61, and rSCC-61 cell lines [176]. A custom k-nearest neighbors script was written to assign greater weight to neighboring genes that have more available data

that can be compared between samples. The distance metric d_{ij} between genes *i* and *j* is calculated using the equation:

$$d_{ij} = \frac{d_{\text{Euclidean},ij}}{n_{ij}^{\alpha}} \tag{19}$$

where $d_{Euclidean,ij}$ is the Euclidean distance between two genes (only for tissues where both genes have values available), n_{ij} is the number of tissues where both genes have values available, and α is a constant with value ≥ 1. The larger the value of α, the greater weight that is given to neighboring genes that have more available data comparable between tissues ($\text{weight}_{ij} \sim 1/d_{ij}$). The k-nearest neighbors script also employs a parameter "minsamples", which is the minimum number of tissues n_{ij} that gene neighbors must both have in order to be valid neighbors. Multiple rounds of k-nearest neighbors is run until no more missing values are imputed. Afterwards, values that are still missing are imputed by the mean value of all genes for a particular tissue. Parameter values used in this paper are: k = 100, minsamples = 3, α = 1.5.

Proteomics data was converted from units of ppm to mmol/gDW through the following conversion:

$$p_{mmol/gDW} = \frac{1000 f_{\text{dw}}}{N_a n_{\text{aa}} m_{\text{aa}} g_{\text{Da}}} \frac{p_{\text{ppm}}}{1000000} \tag{20}$$

where N_a is Avogadro's number, n_{aa} is the average number of amino acids per protein, m_{aa} is the average molecular mass of an amino acid, g_{Da} is the number of grams per Dalton, and p_{dw} is the fraction of the cell's dry weight that is protein. Parameter values used in this paper are n_{aa} = 300, m_{aa} = 100, and f_{dw} = 0.5.

Proteomics data for every gene was used to calculate enzyme concentration values $[E]_j$ using each reaction's gene-reaction rule, which is an and/or statement stating which

genes must be expressed in order for the enzyme catalyzing the particular reaction to be produced. For every tissue and cell type analyzed, gene names were replaced with their respective proteomics value (units of mmol/gDW), "X or Y" statements were replaced with X+Y, and "X and Y" statements were replaced with min(X,Y).

3.2.1.6 Kinetic data

Turnover number data for enzymes within the Recon 2 model were obtained from the BRENDA database API [137]. Figure 3-1 provides an outline of the steps used to get turnover numbers for all reactions with EC numbers available, starting from most accurate to least accurate. When multiple k_{cat} values were available for the correct enzyme and substrate, the mean of these values was used. In order to ensure the physiological accuracy of kinetic parameters used in the model, only BRENDA k_{cat} values measured at a pH between 7-8 were selected, and no k_{cat} values taken from recombinant or mutated enzymes were used.

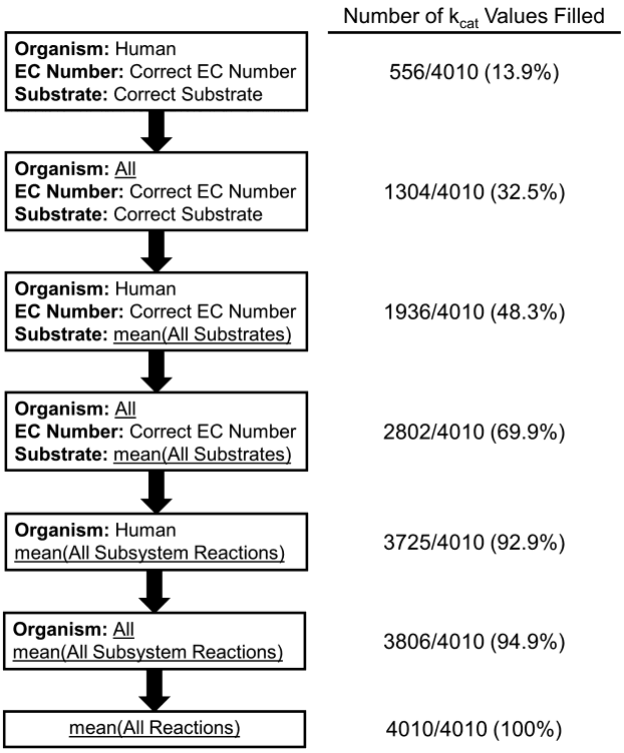


Figure 3-1. Pipeline for determining turnover number (k_{cat}) values for every Recon 2 reaction with an available enzyme commission (EC) number. The number and percentage of needed k_{cat} values filled at every step is given. All data is extracted from the BRENDA database (25). Subsystem refers to the metabolic subsystem assigned to every reaction within the Recon 2 model.

3.2.1.7 Metabolite concentration data

Table 3-1 provides the fixed metabolite concentrations used in the model. Metabolite concentration ranges from Park et al. [177] were incorporated, giving preference to values available in mammalian iBMK cells over yeast cells, and yeast cells over *E. coli* cells (Table 3-2). Human metabolite concentration ranges from The Human Metabolome Database (HMDB) for α-ketoglutarate and pyruvate were used instead of mammalian iBMK cell values, in order to make the IDH1 and ME1 Recon 2 reactions

thermodynamically feasible [178]. Table 3-3 shows the extracellular metabolite concentrations used to model Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) media (ThermoFisher Scientific; 11320033). Table 3-4 shows the experimentally measured cell-type-specific metabolite values and ranges used in SCC-61 and rSCC-61 models [179]. All other metabolite concentrations in the model were allowed to range from e^{-15} to e^{-2} M.

Table 3-1. Compartmental concentrations of water, oxygen, and ions used in the model.

		Metabolite Concentration [M]					
		H_2O	O_2	H^+	Na^+	K^+	Cl^-
Compartment	Extracellular	55	0.006	$10^{-7.4}$	*	*	*
	Cytoplasm	55	0.006	$10^{-7.2}$	0.01	0.16	0.003
	Golgi	55	0.006	$10^{-6.35}$	0.01	0.16	0.003
	Lysosome	55	0.006	$10^{-4.7}$	0.01	0.16	0.003
	Mitochondrion	55	0.006	$10^{-8.0}$	0.01	0.16	0.003
	Nucleus	55	0.006	$10^{-7.2}$	0.01	0.16	0.003
	Endoplasmic Reticulum	55	0.006	$10^{-7.2}$	0.01	0.16	0.003
	Peroxisome	55	0.006	$10^{-7.0}$	0.01	0.16	0.003

Table 3-2. Metabolite concentration upper and lower bounds from Park et al.

Metabolite	Model Name	Lower Bound [M]	Upper Bound [M]
1,3-bisphosphoglycerate	13dpg[c]	6.77E-07	3.28E-06
2-phosphoglycerate	2pg[c]	1.37E-06	9.18E-05
2,3-bisphosphoglycerate	23dpg[c]	1.76E-04	3.16E-04
3-phospho-serine	pser_L[c]	3.45E-04	5.35E-04
3-phosphoglycerate	3pg[c]	2.88E-04	4.63E-04
6-phospho-D-gluconate	6pgc[c]	1.23E-05	2.20E-05
acetoacetyl-CoA	aacoa[c]	1.37E-05	3.47E-05
acetyl-CoA	accoa[c]	2.25E-05	3.50E-05
aconitate	HC00342[c]	8.72E-06	1.34E-05
adenine	ade[c]	1.26E-06	1.71E-06
adenosine	adn[c]	9.44E-08	1.82E-07
adenosine-phosphosulfate	aps[c]	5.10E-06	8.62E-06
ADP	adp[c]	4.52E-04	7.15E-04
ADP-glucose	adpglc[c]	2.83E-06	6.44E-06
α-ketoglutarate	akg[c]	2.90E-04	3.50E-04
alanine	ala_L[c]	6.19E-03	7.77E-03
AMP	amp[c]	3.57E-05	5.02E-05
anthranilate	anth[c]	3.34E-06	3.62E-06
arginine	arg_L[c]	1.75E-04	3.35E-04
asparagine	asn_L[c]	1.56E-04	2.74E-04
aspartate	asp_L[c]	1.38E-02	1.60E-02
ATP	atp[c]	2.40E-03	6.94E-03
carbamoyl-aspartate	cbasp[c]	3.64E-04	9.55E-04
carbon dioxide [c]	co2[c]	6.37E-03	9.55E-03
carbon dioxide [m]	co2[m]	6.37E-03	9.55E-03
citrate	cit[c]	5.59E-04	6.09E-04
citrulline	citr_L[c]	2.41E-02	3.00E-02
CMP	cmp[c]	9.13E-06	1.44E-05
coenzyme-A [c]	coa[c]	8.83E-05	2.12E-02
coenzyme-A [m]	coa[m]	4.00E-03	6.00E-03
CTP	ctp[c]	7.55E-04	1.04E-03
cyclic-AMP	camp[c]	6.29E-08	1.97E-07
cysteine	cys_L[c]	6.19E-05	1.06E-04
cytidine	cytd[c]	1.18E-06	5.67E-06
cytosine	csn[c]	8.65E-06	2.29E-05
dAMP	damp[c]	5.76E-06	2.79E-05
dATP	datp[c]	7.52E-07	1.20E-06
dCDP	dcdp[c]	1.62E-06	2.02E-06
dCMP	dcmp[c]	2.55E-05	4.88E-05
dCTP	dctp[c]	1.96E-06	6.96E-06
deoxyadenosine	dad_2[c]	1.80E-06	4.41E-06
deoxyguanosine	dgsn[c]	4.11E-07	6.62E-07
deoxyribose-5-phosphate	2dr5p[c]	4.82E-05	1.90E-03
dGMP	dgmp[c]	3.91E-05	6.58E-05
dihydroorotate	dhor_S[c]	6.30E-04	8.40E-04
dihydroxyacetonephosphate	dhap[c]	1.47E-03	1.75E-03
dTDP	dtdp[c]	3.37E-04	4.26E-04
dTMP	dtmp[c]	7.20E-06	1.65E-05
dTTP	dttp[c]	4.21E-03	5.08E-03
erythrose-4-phosphate	e4p[c]	7.54E-06	1.58E-05
FAD	fad[c]	3.17E-06	8.03E-06

flavin mononucleotide	fmn[c]	3.84E-05	7.51E-05
fructose-1,6-bisphosphate	fdp[c]	1.40E-03	1.63E-03
fructose-6-phosphate	f6p[c]	8.18E-05	1.27E-04
fumarate [c]	fum[c]	2.83E-04	4.90E-04
fumarate [m]	fum[m]	4.80E-04	4.90E-04
GDP	gdp[c]	1.96E-05	4.08E-05
gluconate	glcn[c]	1.24E-04	2.98E-04
glucosamine-6-phosphate	gam6p[c]	9.59E-04	1.39E-03
glucose-6-phosphate	g6p[c]	6.74E-04	1.05E-03
glutamate	glu_L[c]	4.36E-02	8.39E-02
glutamine	gln_L[c]	1.62E-02	1.82E-02
glutathione	gthrd[c]	2.95E-03	3.22E-03
glutathione disulfide	gthox[c]	1.57E-05	2.03E-05
glyceraldehyde-3-phosphate	g3p[c]	1.28E-04	1.52E-04
glycerate	glyc_R[c]	6.44E-04	3.08E-03
glycine	gly[c]	3.61E-03	3.81E-03
GMP	gmp[c]	1.12E-05	2.49E-05
GTP	gtp[c]	6.52E-04	7.03E-04
guanine	gua[c]	1.32E-04	2.69E-04
guanosine	gsn[c]	1.19E-06	1.51E-06
histidine	his_L[c]	3.86E-04	4.34E-04
homocysteine	hcys_L[c]	3.65E-04	3.75E-04
IDP	idp[c]	1.76E-05	3.22E-05
IMP	imp[c]	9.06E-06	1.55E-05
inosine	ins[c]	5.38E-07	2.11E-06
isocitrate [c]	icit[c]	4.68E-06	4.29E-05
isocitrate [m]	icit[m]	2.10E-05	3.38E-05
isoleucine	ile_L[c]	1.66E-03	1.86E-03
ITP	itp[c]	1.38E-04	3.03E-04
leucine	leu_L[c]	1.66E-03	1.86E-03
lysine	lys_L[c]	4.58E-04	5.54E-04
malate	mal_L[c]	1.06E-03	1.72E-03
malonyl-CoA	malcoa[c]	4.76E-06	5.13E-06
methionine	met_L[c]	6.19E-04	6.59E-04
methylmalonic acid	HC00900[c]	3.76E-05	1.08E-04
myo-inositol	inost[c]	4.22E-06	7.75E-06
N-acetyl-glucosamine-1/6-phosphate	acgam1p[c]	4.38E-06	1.06E-05
N-acetyl-L-alanine	CE1554[c]	8.77E-06	1.19E-05
N-acetyl-L-aspartic acid	Nacasp[c]	2.77E-03	3.03E-03
N-acetyl-ornithine	acorn[c]	1.93E-04	2.65E-04
NAD+	nad[c]	4.72E-04	5.32E-04
NADH	nadh[c]	6.36E-05	8.82E-05
NADP+	nadp[c]	2.55E-05	3.14E-05
NADPH	nadph[c]	5.87E-05	7.21E-05
ornithine	orn[c]	4.04E-03	4.94E-03
orotate	orot[c]	2.19E-06	1.46E-05
oxaloacetate [c]	oaa[c]	2.81E-07	8.55E-07
oxaloacetate [m]	oaa[m]	1.18E-06	3.31E-06
phenylalanine	phe_L[c]	7.97E-04	8.84E-04
phenylpyruvate	phpyr[c]	1.71E-03	1.82E-03
phosphate	pi[c]	4.00E-03	6.00E-03
phosphoenolpyruvate	pep[c]	6.86E-06	1.98E-05

proline	pro_L[c]	1.13E-03	1.34E-03
propionyl-CoA	ppcoa[c]	3.88E-06	7.29E-06
PRPP	prpp[c]	3.67E-05	5.73E-05
pyruvate	pyr[c]	2.70E-05	3.70E-04
quinolinate	quln[c]	2.41E-06	5.49E-05
riboflavin	ribflv[c]	1.72E-05	2.11E-05
ribose-5-phosphate	r5p[c]	2.84E-05	4.01E-05
ribulose-5-phosphate	ru5p_D[c]	4.10E-06	5.85E-06
S-adenosyl-L-methionine	amet[c]	1.19E-04	2.84E-04
sedoheptulose-7-phosphate	s7p[c]	1.28E-05	1.81E-05
serine	ser_L[c]	4.75E-03	4.97E-03
sn-glucerol 3-phosphate	glyc3p[c]	2.54E-04	3.08E-04
succinate	succ[c]	2.87E-04	4.30E-04
succinyl-CoA [c]	succoa[c]	1.42E-04	3.83E-04
succinyl-CoA [m]	succoa[m]	3.49E-07	3.16E-03
taurine	taur[c]	5.21E-06	6.06E-06
threonine	thr_L[c]	6.44E-03	6.94E-03
thymidine	thymd[c]	1.88E-06	3.40E-06
trehalose	tre[c]	6.53E-03	1.03E-02
tryptophan	trp_L[c]	1.72E-04	1.89E-04
tyrosine	tyr_L[c]	8.88E-04	9.88E-04
UDP	udp[c]	1.25E-04	1.42E-04
UDP-glucose	udpg[c]	1.28E-03	1.77E-03
UDP-glucuronate	udpglcur[c]	8.45E-05	1.11E-04
UDP-N-acetyl-glucosamine	uacgam[c]	8.80E-03	9.14E-03
UMP	ump[c]	1.33E-05	1.57E-05
uridine	uri[c]	1.96E-03	2.24E-03
UTP	utp[c]	1.55E-03	1.98E-03
valine	val_L[c]	1.44E-03	1.58E-03
xylulose-5-phosphate	xu5p_D[c]	2.34E-05	3.31E-05

Table 3-3. Metabolite formulation of DMEM/F-12 media used to model the extracellular environment.

Metabolite	Model Name	Concentration [M]
alanine	ala_L[e]	5.00E-05
arginine	arg_L[e]	6.99E-04
asparagine	asn_L[e]	5.00E-05
aspartate	asp_L[e]	5.00E-05
biotin	btn[e]	1.43E-08
Ca2+	ca2[e]	1.06E-03
choline	chol[e]	6.41E-05
Cl-	cl[e]	1.29E-01
cysteine	cys_L[e]	9.98E-05
cystine	Lcystin[e]	1.00E-04
Fe2+	fe2[e]	1.50E-06
Fe3+	fe3[e]	1.24E-07
folate	fol[e]	6.01E-06
glucose	glc_D[e]	1.75E-02
glutamate	glu_L[e]	5.00E-05
glutamine	gln_L[e]	2.50E-03
glycine	gly[e]	2.50E-04
HCO3-	hco3[e]	2.90E-02
histidine	his_L[e]	1.50E-04
hypoxanthine	hxan[e]	1.50E-05
inositol	inost[e]	7.00E-05
isoleucine	ile_L[e]	4.16E-04
K+	k[e]	4.16E-03
leucine	leu_L[e]	4.51E-04
linoleate	lnlc[e]	1.50E-07
lipoate	lipoate[e]	5.10E-07
lysine	lys_L[e]	4.99E-04
methionine	met_L[e]	1.16E-04
Na+	na1[e]	1.52E-01
niacinamide	ncam[e]	1.66E-05
panthothenate	pnto_R[e]	9.39E-06
phenylalanine	phe_L[e]	2.15E-04
phosphate	pi[e]	9.53E-04
proline	pro_L[e]	1.50E-04
putrescine	ptrc[e]	5.03E-07
pyridoxine	pydxn[e]	9.77E-06
pyruvate	pyr[e]	5.00E-04
riboflavin	ribflv[e]	5.82E-07
serine	ser_L[e]	2.50E-04
sulfate	so4[e]	4.10E-04
thiamine	thm[e]	6.44E-06
threonine	thr_L[e]	4.49E-04
thymidine	thymd[e]	1.51E-06
tryptophan	trp_L[e]	4.42E-05
tyrosine	tyr_L[e]	2.14E-04
valine	val_L[e]	4.52E-04
vitamin B12	aqcobal[e]	5.02E-07
Zn2+	HC02172[e]	1.50E-06

Table 3-4. Experimentally-measured metabolite concentration values and ranges in SCC-61 and rSCC-61 cells.

Metabolite	Model Name	SCC-61			rSCC-61		
		Concentration [M]	Lower Bound [M]	Upper Bound [M]	Concentration [M]	Lower Bound [M]	Upper Bound [M]
aspartate	asp_L[c]	-	4.94E-06	9.19E-06	-	6.81E-07	2.32E-06
cholesterol	chsterol[c]	7.38E-02	-	-	3.44E-02	-	-
cysteine	cys_L[c]	-	9.88E-07	1.48E-06	-	1.29E-04	2.37E-04
glutamate	glu_L[c]	-	5.79E-05	9.97E-05	-	1.06E-05	1.55E-05
glutathione	gthrd[c]	-	1.90E-05	2.57E-05	-	8.75E-06	1.00E-05
glutathione disulfide	gthox[c]	-	3.78E-08	6.04E-08	-	1.52E-08	3.35E-08
hexadecanoate (C16:0)	hdca[c]	6.58E-04	-	-	1.47E-03	-	-
hexadecenoate (C16:1)	hdcea[c]	1.90E-04	-	-	4.95E-04	-	-
octadecadienoate (C18:2)	ocdcya[c]	9.75E-05	-	-	9.00E-05	-	-
octadecanoate (C18:0)	ocdca[c]	3.45E-04	-	-	3.75E-04	-	-
octadecenoate (C18:1)	ocdcea[c]	1.05E-03	-	-	1.48E-03	-	-
phosphatidylcholine	pchol_hs[c]	6.41E+00	-	-	6.98E+00	-	-
phosphatidylethanolamine	pe_hs[c]	2.52E+00	-	-	2.96E+00	-	-
phosphatidylglycerol	pglyc_hs[c]	8.90E-03	-	-	5.64E-02	-	-
phosphatidylserine	ps_hs[c]	1.00E+00	-	-	1.48E+00	-	-
sphingomyelin	sphmyln_hs[c]	8.48E-01	-	-	7.93E-01	-	-
tetradecanoate (C14:0)	ttdca[c]	5.75E-05	-	-	2.50E-04	-	-

3.2.1.8 Random sampling

Reaction flux distributions for SCC-61 and rSCC-61 cell types were obtained by performing random sampling of the solution space for Equation 7 with all thermodynamically-infeasible reactions removed from the model. optReduceModel and optGpSampler were used to perform random sampling, and only solutions that achieved at least 95% of the maximum objective function value were kept [133].

3.2.1.9 Modeling gene knockdown

siRNA gene knockdown was modeled by decreasing the protein expression of individual genes by 75% (predicted percent knockdown using N-TER Nanoparticle siRNA Transfection System provided by the manufacturer is ≥ 70%). These decreased protein expression values were used to calculate enzyme concentration values $[E]_j$ using each reaction's gene-reaction rule, as detailed above.

3.2.2 Experimental methods

3.2.2.1 siRNA screen

SCC-61 and rSCC-61 cells were maintained in DMEM/F12 medium that was supplemented with 10% fetal bovine serum (Sigma-Aldrich, F4135) at 37°C and 5% CO2. For RNAi experiments, cells were plated on 96-well plates at seeding densities of 4 x 10^3 cells per well. White opaque-bottom 96-well plates were used for the cell viability and NADPH luminescent assays. 24 hours after plating, cells were transfected using the N-TER Nanoparticle siRNA Transfection System (Sigma-Aldrich, N2913) and the MISSION siRNA Oxidoreductase Gene Panel (Sigma-Aldrich, SI12100). *Drosophila* sequences were used for transfection controls as provided by the manufacturer. Three siRNA sequences for each of 338 gene targets in the library were pooled, and wells were each treated with a final siRNA concentration of 40nM. All transfections and controls were performed in triplicate. Serum-containing transfection medium was used.

3.2.2.2 Cell viability assay

4 hours after transfection, β-lapachone diluted in serum-containing media was added to each well at a final concentration of 3μM (ID50 of rSCC-61 cells). After 24 hours of β-lapachone exposure, the CellTiter-Glo Luminescent Cell Viability Assay was performed (Promega, G7572) according to manufacturer. Luminescence readings were performed using a Bio-Tek Synergy 4 multimode reader.

3.2.2.3 NADPH assay

Plates with transfected cells were split into control and treatment plates. 24 hours after transfection, β-lapachone diluted in serum-containing media was added to the treatment plates at a final concentration of 3 μM. An equal volume of media was added to

the control plates. After 2 hours of β-lapachone exposure, the NADP/NADPH-Glo Assay (Promega, G9081) was performed on both control and treatment plates according to manufacturer. Luminescence readings were performed using a Bio-Tek Synergy 4 hybrid microplate reader.

3.3 Results

3.3.1 Construction of genome-scale metabolic models for HNSCC cell lines

The cell lines under investigation are a parental line (SCC-61) and a radiation-resistant line selected after successive survival after ionizing radiation (rSCC-61); these cell lines have been extensively characterized for metabolic, transcriptomic, proteomic, and epigenetic changes [171, 175, 179]. We hypothesized that β-lapachone sensitivity of gene silencing in SCC-61 and rSCC-61 cells could be predicted by FBA models of NAD(P)H production. To test this hypothesis, a key component of our model framework is the use of a novel objective function that optimized the yield of $NAD(P)^+ \rightarrow NAD(P)H$ conversion while maintaining sufficient ATP production to obtain biologically viable results; this is a departure from prior FBA models of cancers that solely optimize biomass or ATP production.

Our model comprises three main components (Figure 3-2). The first component is the core metabolic network, which is based on the Recon 2 genome-scale human metabolic network with 5063 metabolites and 7440 reactions [135]. The core network represents the collection of all known human metabolites and associated reactions and is the starting point to create cell-type-specific models. We have made additional changes to the Recon 2 metabolic network to address missing or inaccurate redox-based reaction information. The

second component of the model is the incorporation of an upper bound v_{max} to every reaction flux within the network. This upper bound was set as the product of the turnover number k_{cat} (units of s^{-1}) and abundance [E] (units of mmol/gDW) of the enzyme that catalyzes each particular reaction; this product is equivalent to the Michaelis–Menten constant V_{max}. If a reaction within the model is not enzyme-catalyzed, an arbitrarily large upper bound is given. Turnover numbers for every enzyme in the Recon 2 model were obtained from the BRENDA database [137] (Figure 3-1), and enzyme abundance values for cell types of interest were obtained using proteomic and transcriptomic data from multiple sources. The last component of the model is the incorporation of thermodynamic constraints on each reaction [12]. If the ΔG of any reaction could not achieve a value <0, then thermodynamically it cannot carry a net flux, and therefore the upper flux bound v_{max} was set to zero. Standard reaction free energies used to calculate ΔG were obtained from the Virtual Metabolic Human database [135]. Metabolite concentrations used include experimentally verified cellular metabolite concentration ranges, components of cell culture media to simulate the extracellular environment, and measured concentration values specific to the cell types of interest. The incorporation of enzyme abundances and metabolite concentrations allows for the creation of cell-type-specific models, and incorporation of enzyme turnover numbers and reaction standard free energies further improves the biological accuracy of model predictions.

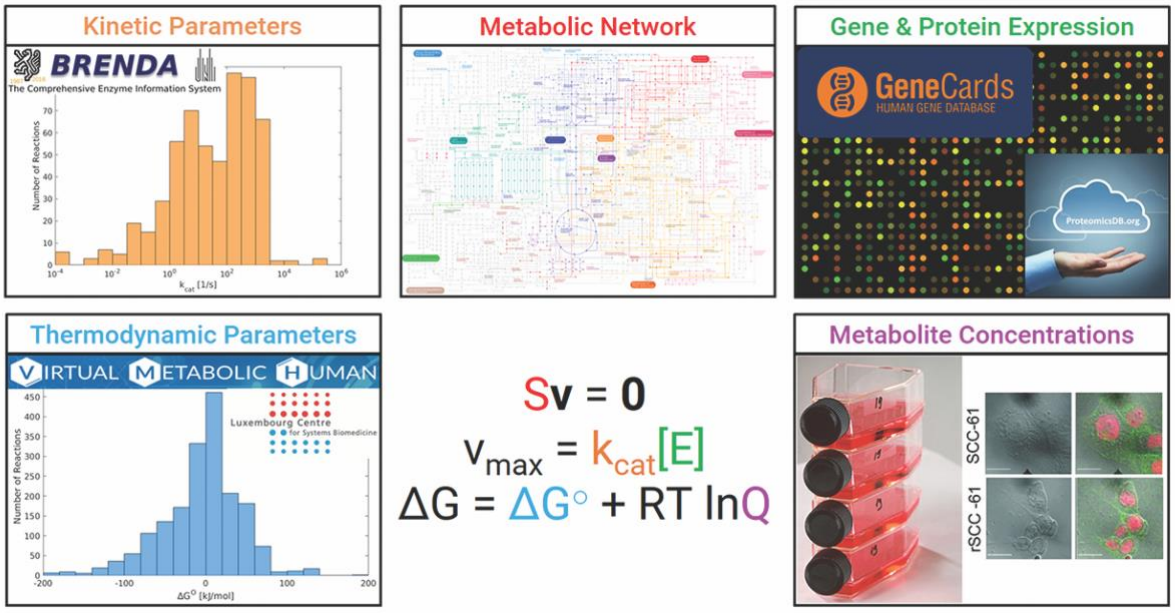


Figure 3-2. Incorporation of disparate sources of biological information into the mathematical modeling framework.

Models specific to SCC-61 and rSCC-61 cell types were constructed to better understand their differences in NAD(P)H production. Because proteome-wide enzyme abundance values for these two cell types were not available, these values were estimated from SCC-61 and rSCC-61 gene expression data by extrapolating onto proteomic values measured in CAL-27 head and neck cancer cells; missing values where no proteomic data was available were imputed. This methodology conserves most of the correlation between experimental gene expression data (used to estimate enzyme abundance) and the small amount of experimental proteomic data available for SCC-61 and rSCC-61 cells (Figure 3-3, Figure 3-4).

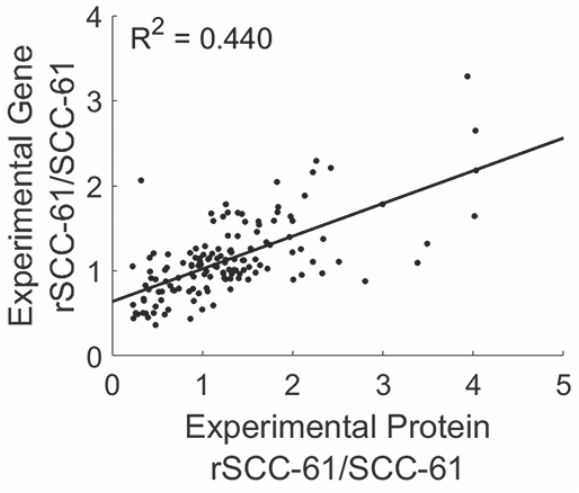


Figure 3-3. Correlation between experimental proteomic and experimental gene expression values (as ratios of rSCC-61 to SCC-61) corresponding to the same NCBI gene ID.

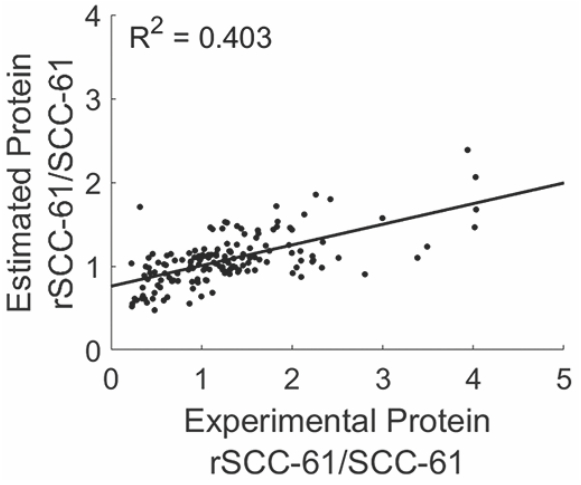


Figure 3-4. Correlation between experimental proteomic and model proteomic values. Estimation of model proteomic values from experimental gene expression

data conserves most of the original correlation between experimental proteomic and gene expression values shown in Figure 3-3.

To determine whether the incorporation of thermodynamic constraints is necessary to achieve accurate model predictions, the net conversion of $NADP^+$ to NADPH was compared in both cell types, with and without standard free energy values (ΔG_o) and experimentally measured metabolite values (Q) included in the model (Figure 3-5). Only when both data sets were included was the conversion of $NADP^+$ to NADPH greater in rSCC-61 cells than in SCC-61 cells. In this case, the ratio of net conversions (rSCC-61/SCC-61 = 1.40) was very close to the experimentally measured ratio of NADPH/$NADP^+$ concentration between rSCC-61 cells and SCC-61 cells (rSCC-61/SCC-61 = 1.56) [179].

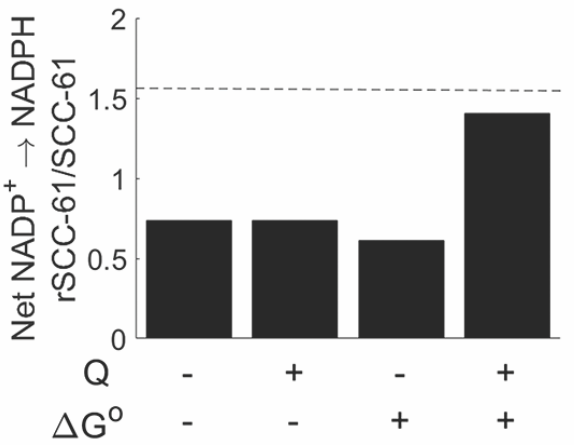


Figure 3-5. Net conversion of $NADP^+$ to NADPH, expressed as a ratio of rSCC-61/SCC-61. Values obtained from the model with and without utilizing metabolite concentration data (Q) and reaction standard free energy values (ΔG_o) are shown.

The *dotted line* corresponds to the experimentally measured ratio of NADPH/$NADP^+$ concentration between rSCC-61 and SCC-61 cells.

3.3.2 Predicted NADPH production across the HNSCC metabolome

The flux distributions of major NADPH-producing reactions shown in Figure 1 were calculated in both SCC-61 and rSCC-61 cell types (Figure 3-6). Flux values span multiple orders of magnitude and most predicted values are within one order of magnitude of experimentally measured flux values determined using ^{13}C-tracers in other cell types [177]. Even though the upper flux bound on the glutamate dehydrogenase GLUD1/GLUD2 reaction is higher in SCC-61 cells than rSCC-61 cells (due to greater gene expression), the observed flux values through this reaction in rSCC-61 cells are significantly greater. While rSCC-61 cells exhibit a flux distribution constrained relatively close to the upper flux bound, SCC-61 cell flux deviates greatly from the GLUD1/GLUD2 upper bound. This result highlights the importance of developing genome-scale models for accurate prediction of metabolic phenotypes instead of simply comparing individual gene or protein expression values. While a large difference in flux through the GLUD1/GLUD2 reaction was found between SCC-61 and rSCC-61 cell types, most reactions did not display substantial differences in flux distribution. One exception is nicotinamide nucleotide transhydrogenase (NNT), which shows significantly greater flux in rSCC-61 cells than SCC-61 and was previously found to be a potential prognostic marker in HNSCC [180]. However, because NNT flux values are several orders of magnitude lower than flux values through other NADPH-producing reactions, differences in NNT flux between SCC-61 and rSCC-61 cell types may not be biologically significant. Note that no nonzero solutions for the malic enzyme ME2/ME3 reaction were obtained in SCC-61 cells; this artifact is likely

due to the unrestricted flux of malate and pyruvate across the mitochondrial membrane to the cytoplasm, along with the ability of ME1 to catalyze the same reaction as ME2/ME3 in the cytoplasm.

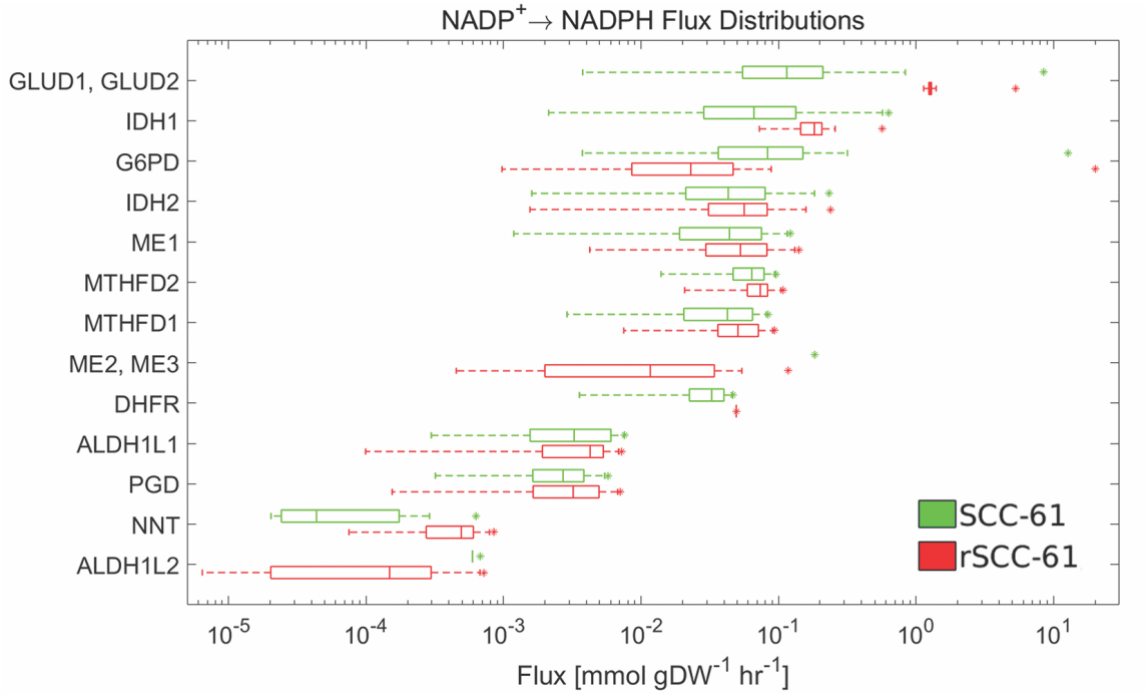


Figure 3-6. Box plot showing the flux distributions in both SCC-61 and rSCC-61 cells of major NADPH biogenesis reactions. Note that no nonzero solutions for the ME2/ME3 reaction were obtained in SCC-61 cells (see text). *Boxes* represent the following percentiles: 2.5th, 25th, 50th, 75th, and 97.5th. *Dots* represent the upper flux bound for each reaction ($v_{max} = k_{cat}[E]$).

To determine whether the calculated flux through NADPH-producing reactions correlates with the importance of the corresponding gene's expression toward total cellular NADPH production, the maximum net $NADP_+ \rightarrow$ NADPH flux in both SCC-61 and rSCC-61 cell types was calculated after 75% simulated gene knockdown (Figure 3-7). Surprisingly, even though the GLUD1/GLUD2 reaction had the largest calculated flux

values among all reactions in both cell types, knockdown of these genes caused very little change in total cellular NADPH production. However, some gene perturbations had a substantial impact on NADPH production, differentially impacting rSCC-61 cells (isocitrate dehydrogenase IDH1, G6PD, ME2) or SCC-61 cells (IDH2). These differences between reaction flux distributions and impact of gene knockdown suggest that the NADPH-producing reactions with the largest flux are not necessarily the most important toward global NADPH production. Instead, metabolic flux can be diverted to other parts of the metabolic network to maintain NADPH production in these cases.

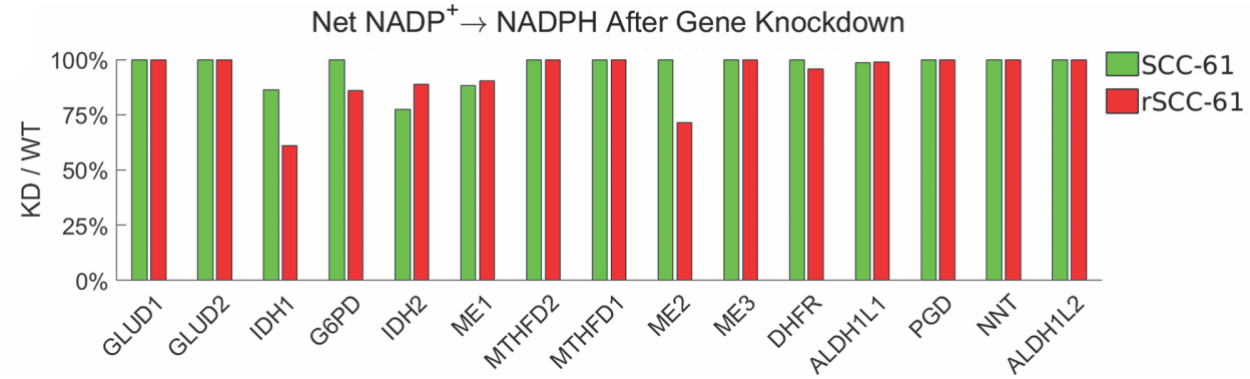


Figure 3-7. Effect of 75% simulated gene knockdown on net conversion of $NADP^+$ to NADPH in both SCC-61 and rSCC-61 cells, expressed as a percentage of net conversion before knockdown.

3.3.3 Simulated silencing of IDH1 highlights NADPH flux re-routing

As shown in Figure 3-7, simulated knockdown of a gene involved in NADPH production does not necessarily impact the total cellular production of NADPH. Instead, the cell can reroute flux toward other pathways to achieve the goal of maximizing NADPH production. To further investigate this rerouting of flux, the flux distributions of SCC-61

and rSCC-61 cells after simulated 75% knockdown of five canonical NADPH-producing genes from Figure 3-6, as well as methylenetetrahydrofolate reductase (MTHFR), were calculated (Figure 3-8, Figure 3-9, Figure 3-10). The condition generating the most interesting insight in flux redistribution was the simulated silencing of IDH1, which increased flux through GLUD1/GLUD2 toward its v_{max} value in the SCC-61 line and decreased flux through GLUD1/GLUD2 in the rSCC-61 cell line. A likely explanation of this differential effect is that both IDH1 and GLUD1/GLUD2 reactions produce α-ketoglutarate. Because the objective function is to maximize conversion of $NADP^+$ to NADPH, the preferred direction of both of these reactions is to produce α-ketoglutarate and NADPH. Due to other model constraints on α-ketoglutarate consumption, there is a limit to how much combined flux can occur through IDH1 and GLUD1/GLUD2. Therefore, reduction in protein levels of IDH1 affords the opportunity of cells to maintain NADPH production by rerouting flux through GLUD1/GLUD2. However, whereas SCC-61 cells normally operate far from the maximum GLUD1/GLUD2 flux bound and can therefore increase this flux, rSCC-61 cells cannot increase GLUD1/GLUD2 flux due to other constraints on the genome-wide metabolic network. A similar result is obtained with simulated G6PD silencing, accounting for the greater impact of G6PD knockdown on NADPH production in rSCC-61 cells than SCC-61 cells (Figure 3-9, Figure 3-10).

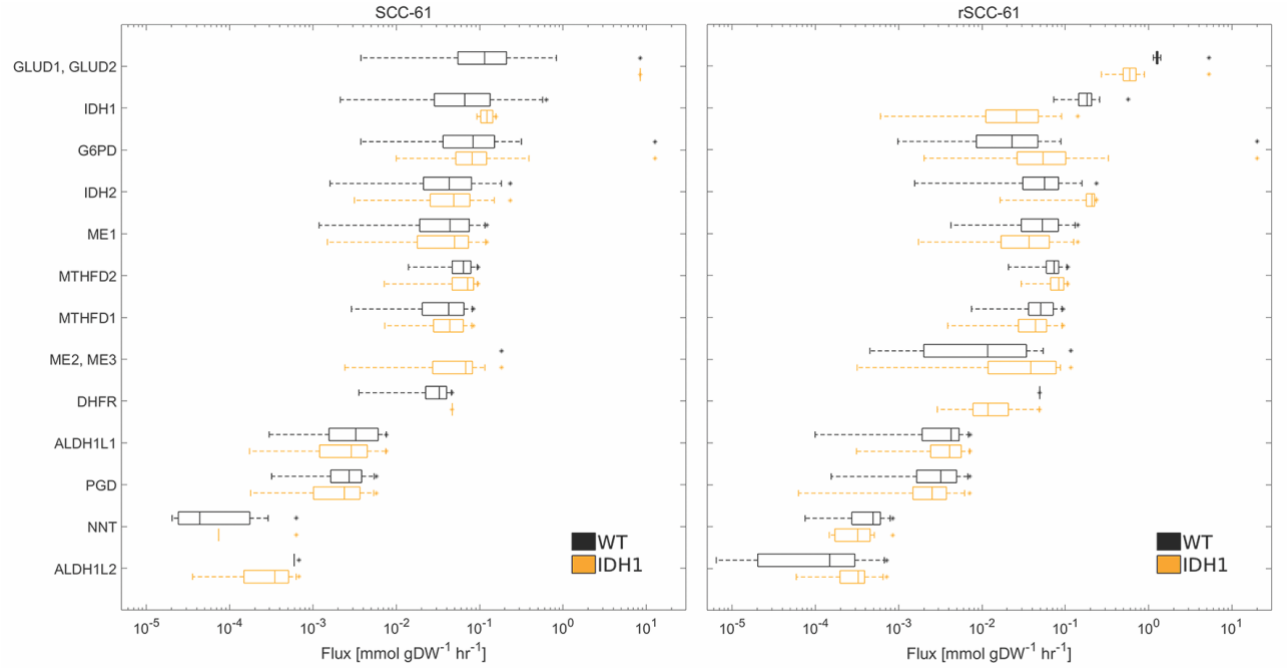


Figure 3-8. Alternative NADPH production pathways after knockdown of IDH1. Flux distributions of simulated wild-type and 75% IDH1 knockdown SCC-61 and rSCC-61 cells. *Boxes* represent the following percentiles: 2.5th, 25th, 50th, 75th, and 97.5th. *Dots* represent the upper flux bound for each reaction ($v_{max} = k_{cat}[E]$), taking into account the lowered IDH1 flux bound in IDH1 knockdown cells.

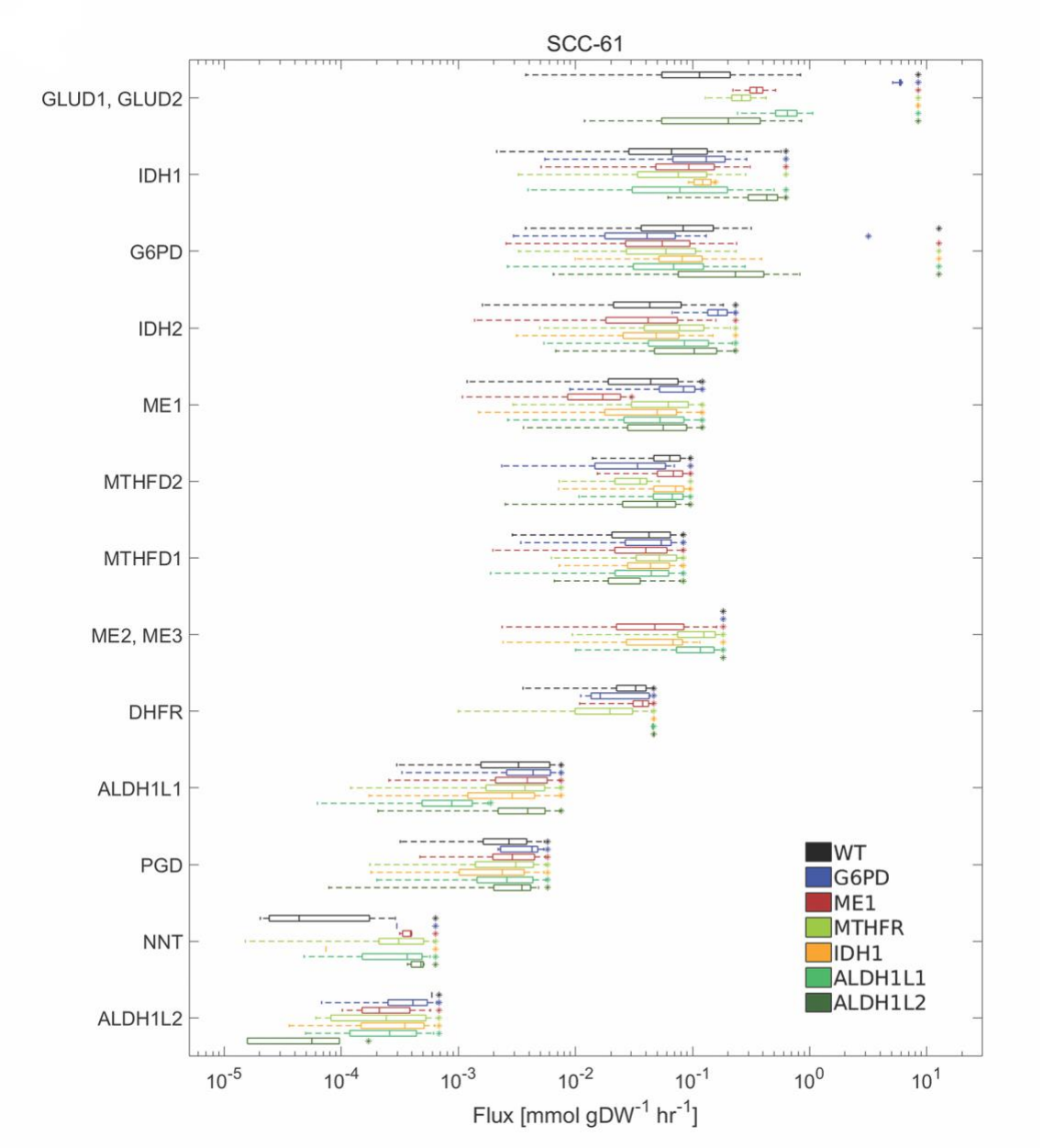


Figure 3-9. Alternative NADPH production pathways after 75% knockdown of major NADPH-biogenesis genes, as well as MTHFR, in SCC-61 cells.

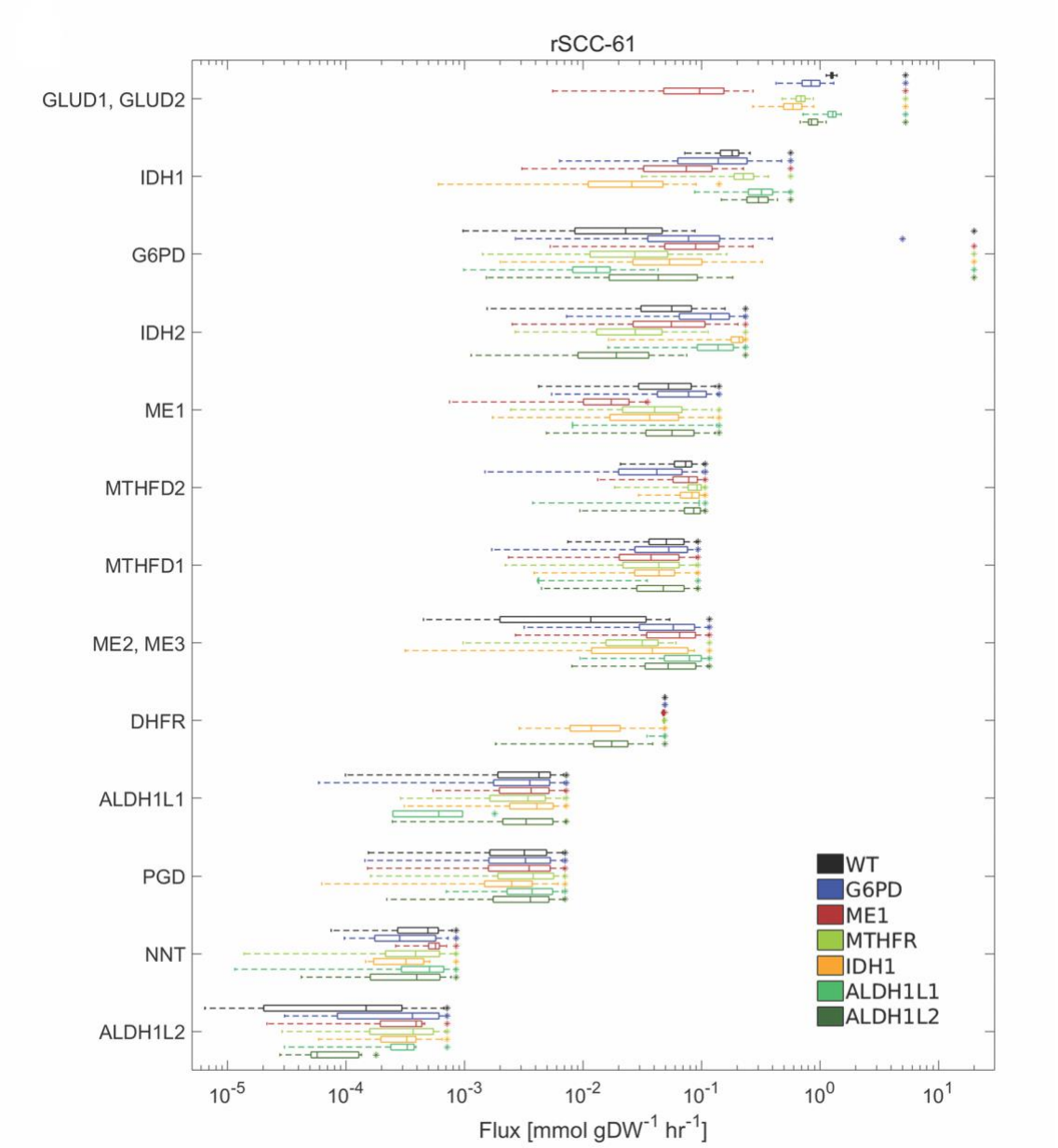


Figure 3-10. Alternative NADPH production pathways after 75% knockdown of major NADPH-biogenesis genes, as well as MTHFR, in rSCC-61 cells.

3.3.4 Upregulation of NADH-producing fluxes in rSCC-61

Using an objective function of maximizing NADH production, parsimonious flux balance analysis was used to obtain a single representative flux vector in both SCC-61 and rSCC-61 cell lines that maximizes both the production of NAD^+, as well as reduction of

NAD^+ to NADH. The rSCC-61 model showed greater total NADH production than the SCC-61 model, with rSCC-61/SCC-61 = 1.24 (Figure 3-11). Both models displayed significant fluxes through the NAD^+ salvage, purine salvage, and pentose phosphate pathways. The pentose phosphate pathway produces ribose 5-phosphate (R5P), which is converted into adenosine monophosphate and phosphoribosyl pyrophosphate by phosphoribosyl pyrophosphate synthase (PRPS). Phosphoribosyl pyrophosphate is converted to nicotinamide mononucleotide and NAD^+ by NAMPT and NMNAT, respectively, in the NAD^+ salvage pathway, supplying both cell lines with sufficient amounts of NAD^+. Adenosine monophosphate is converted to other purines and members of the purine salvage pathway, eventually being excreted as uric acid. Through this pathway, NAD^+ produced from the NAD^+ salvage pathway can be reduced to NADH by both inosine monophosphate dehydrogenase and xanthine oxidase. Additionally, ribose 1-phosphate (R1P) produced from xanthosine can be reconverted into R5P, entering back into the cycle. Thus, the interconnection of the pentose phosphate pathway, NAD^+ salvage pathway, and purine salvage pathway plays an important role in both NAD^+ synthesis and reduction to NADH in both cell lines.

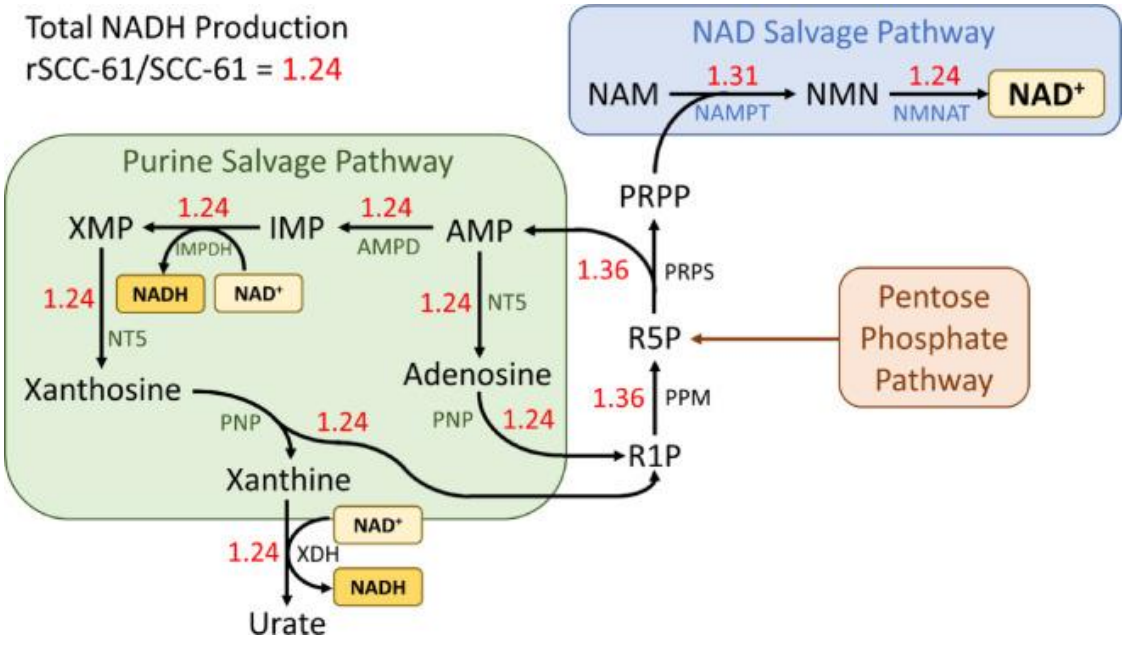


Figure 3-11. Parsimonious flux balance analysis results in SCC-61 and rSCC-61 cell lines. Cytosolic NADH production was maximized to determine the most pertinent metabolic pathways and reactions toward both NAD_+ production and reduction of NAD_+ to NADH. Red numbers indicate the ratio of predicted flux values for each reaction between rSCC-61/SCC-61 cell lines.

To find selective targets of NAD_+ and NADH production in radiation-resistant tumors, the metabolic reactions that had significantly greater flux in the radiation-resistant rSCC-61 cell line compared to the radiation-sensitive SCC-61 cell line were determined, such that the rSCC-61/SCC-61 flux ratio was larger than the ratio of total NADH production (rSCC-61/SCC-61 = 1.24). Both phosphopentomutase (PPM) and PRPS had flux ratios of 1.36, and NAMPT had a flux ratio of 1.31. Previous studies have shown that inhibition of NAMPT decreases NAD_+/NADH levels, increases levels of ROS, and sensitizes cancer cells to β-lapachone treatment [170, 181]. These computational and

experimental findings suggest that combining NAMPT inhibitors with β-lapachone and/or ionizing radiation may be an effective treatment strategy for radiation-resistant tumors.

3.3.5 ß-lapachone sensitivity and effects on NADPH production

Through its redox cycling mechanism with NQO1, β-lapachone treatment results in a net conversion of NADPH to $NADP_+$, thus lowering the NADPH/$NADP_+$ concentration ratio. Subsequent hyperactivation of PARP1 by DNA damage and elevated pools of NAD_+ confers loss of NAD_+ in exposed cells. With cell line-specific models of NADPH production established for SCC-61 and rSCC-61 that account for experimentally observed differences in NADPH/$NADP_+$, we next asked if these phenotypes correspond to differences in each line's ability to redox cycle β-lapachone. To determine what effect siRNA knockdown of canonical NADPH-producing genes as well as MTHFR has on changes in NADPH/$NADP_+$ due to redox cycling, we measured relative changes in [NADPH] and [$NADP_+$] in each knockdown in the presence of 3 μM β-lapachone for 2 h. To quantify the effect of β-lapachone on NADPH/$NADP_+$, a β-lapachone effect parameter was developed:

$$\beta\text{-Lapachone Effect} = \frac{\left(\frac{\text{NADPH}}{\text{NADP}^+}\ \text{w/o β-Lapachone}\right) - \left(\frac{\text{NADPH}}{\text{NADP}^+}\ \text{w/ β-Lapachone}\right)}{\left(\frac{\text{NADPH}}{\text{NADP}^+}\ \text{w/o β-Lapachone}\right)} \quad (21)$$

This effect parameter ranges in value [0,1], where a value of 0 signifies that β-lapachone has no effect on the NADPH/$NADP_+$ratio, and a value of 1 signifies that β-lapachone completely eliminates cellular stores of NADPH. A β-lapachone effect ratio is used to compare the β-lapachone effects between rSCC-61 and SCC-61 cells:

$$\text{β-Lapachone Effect Ratio} = \frac{\text{β-Lapachone Effect on rSCC-61}}{\text{β-Lapachone Effect on SCC-61}} \quad (22)$$

As expected, the effect ratio for Drosophila siRNA control cells is >1 since β-lapachone is known to have a greater effect on rSCC-61 cells than SCC-61 cells (Figure 3-12). ME1, IDH1, and MTHFR knockdowns display substantial increases in effect ratio over control cells, indicating that these knockdowns have a greater negative impact on NADPH/NADP+ ratio in rSCC-61 cells than SCC-61 cells. The G6PD siRNA resulted in a minimal change from control for both cell lines. This finding is consistent with reports of quinone-based chemotherapeutic tolerance for G6PD-deficient individuals [182, 183].

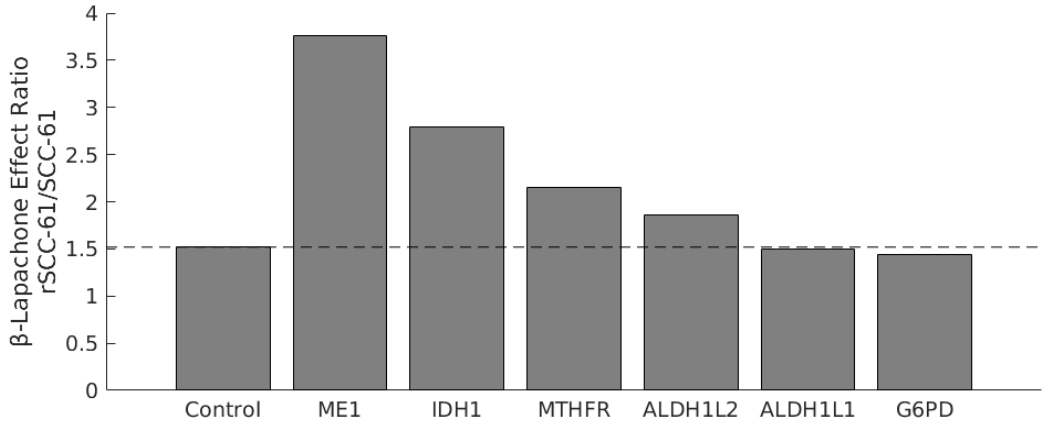


Figure 3-12. β-Lapachone effect ratio in wild-type and siRNA gene knockdown cells. Equations 21 and 22 are used to calculate β-lapachone effect ratio. NADPH/NADP+ ratios were measured both with and without 2 h of exposure of 3 μ*M* β-lapachone. The *dotted line* represents the effect ratio for *Drosophila* control siRNA-transfected cells.

3.3.6 Cell line-specific changes in ß-lapachone sensitivity

Because NQO1-driven quinone redox cycling is a major driver of PARP1 hyperactivation and programmed necrosis, we hypothesized that large changes in the abundance of major NADPH-producing enzymes would directly impact cell viability in the presence of β-lapachone. The rSCC-61 cells have NQO1 mRNA that is over twofold higher than SCC-61 cells. Correspondingly, the measured β-lapachone IC_{50} of rSCC-61 cells by 7-day clonogenic assay is 3 μM, while SCC-61 cells show no sensitivity to the drug. Twenty-four-hour cell viability measurements also indicated sensitivity to 3 μM β-lapachone that is abrogated in the presence of 50 μM dicoumarol, suggesting that the differences in viability are due to NQO1-mediated redox cycling (Figure 3-13, Figure 3-14). To investigate whether our genome-wide estimates of NADPH production can be used to predict modulation of β-lapachone sensitivity in SCC-61 and rSCC-61 cells, model simulations were run in conjunction with an experimental siRNA screen of 338 oxidoreductase genes, 229 of which are found in Recon 2 (Figure 3-15). RNA interference was simulated by individually knocking down gene expression of all 229 genes by 75% and determining the fraction of total cellular NADPH production after knockdown compared with before. Model knockdown of 80 genes caused a greater decrease in NADPH production in rSCC-61 cells than SCC-61 cells, compared with 56 gene knockdowns causing a greater decrease in SCC-61 cells than rSCC-61 cells (Figure 3-16). Experimental knockdown of 59 genes decreased cell viability in rSCC-61 cells, while knockdown of only 10 genes in SCC-61 cells decreased viability in the presence of β-lapachone (Figure 3-17). In addition, experimental knockdown of 181 genes had a more negative impact on cell viability upon β-lapachone exposure in rSCC-61 cells than SCC-61 cells, compared with

48 gene knockdowns having a more negative impact in SCC-61 cells than rSCC-61 cells. This corroboration between model and experimental results suggests that the SCC-61 metabolic network's response to β-lapachone is more robust than in rSCC-61 cells, likely due to SCC-61's ability to reroute NADPH-producing flux through alternative metabolic pathways.

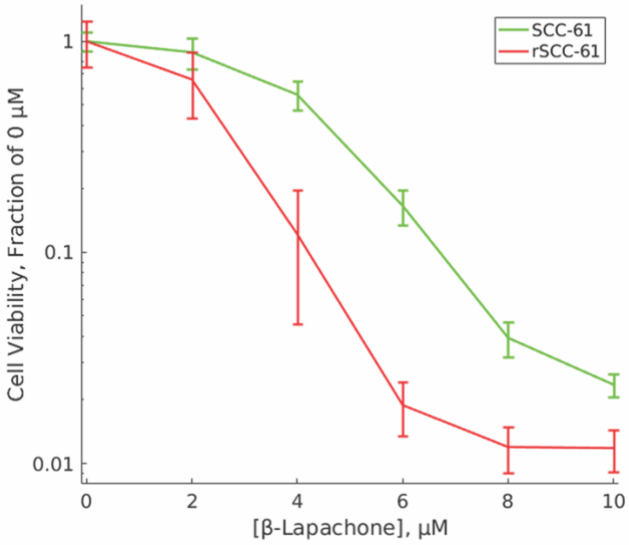


Figure 3-13. Cell viability of SCC-61 and rSCC-61 cells at varying levels of ß-lapachone concentration, expressed as a fraction of cell viability at 0 µM ß-lapachone. Error bars represent ± 1 standard deviation.

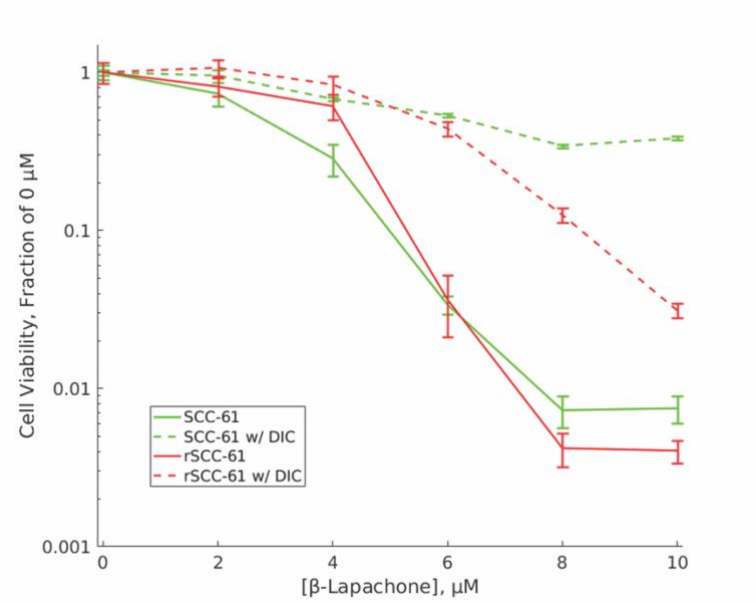


Figure 3-14. Cell viability of SCC-61 and rSCC-61 cells at varying levels of ß-lapachone concentration, with and without treatment with 50 µM dicoumarol, expressed as a fraction of cell viability at 0 µM ß-lapachone. Error bars represent ± 1 standard deviation.

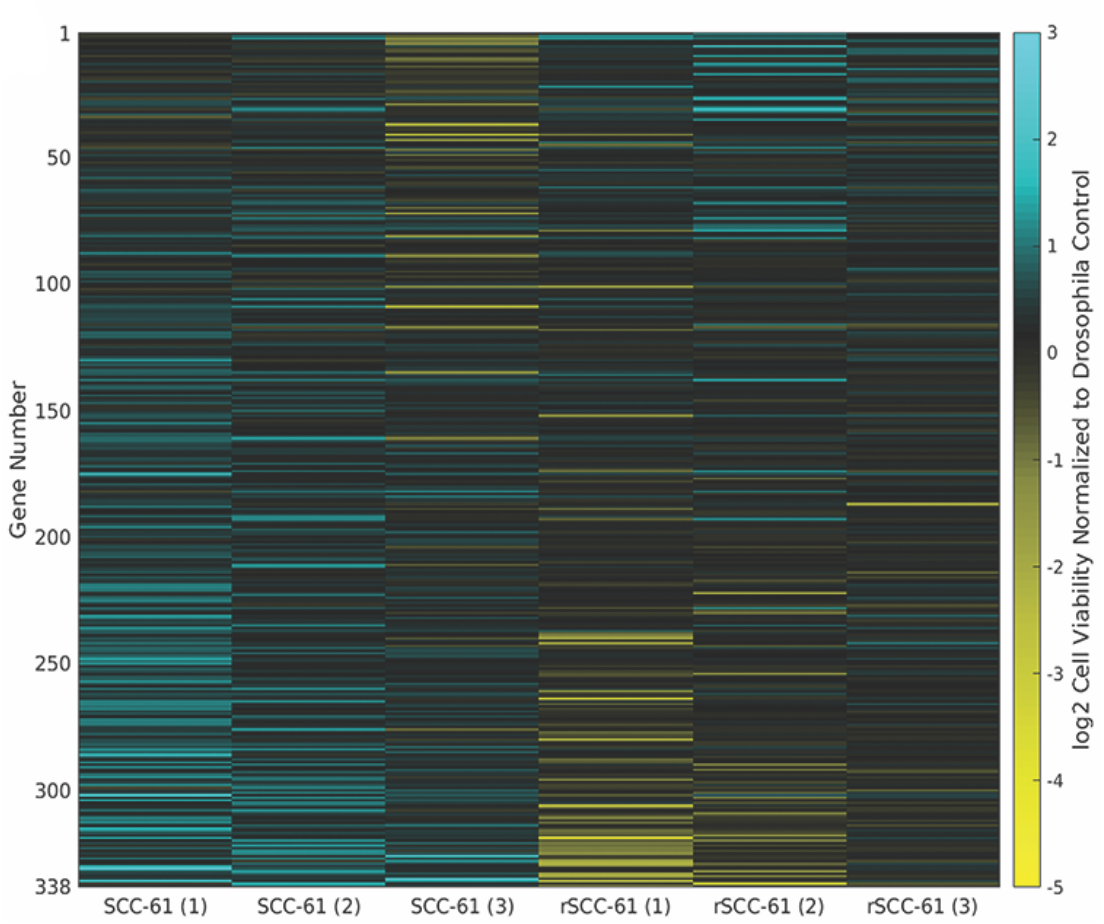


Figure 3-15. Heat map showing cell viability of all 338 gene knockdowns after 24 h of treatment with 3 µ*M* β-lapachone in three replicate experiments. Cell viability values are normalized to those from *Drosophila* control siRNA-transfected cells, and the log2 of normalized values was taken (values >0 signify increased cell viability after knockdown, values <0 signify decreased cell viability after knockdown). Genes are ordered from highest to lowest rSCC-61/SCC-61 cell viability ratio.

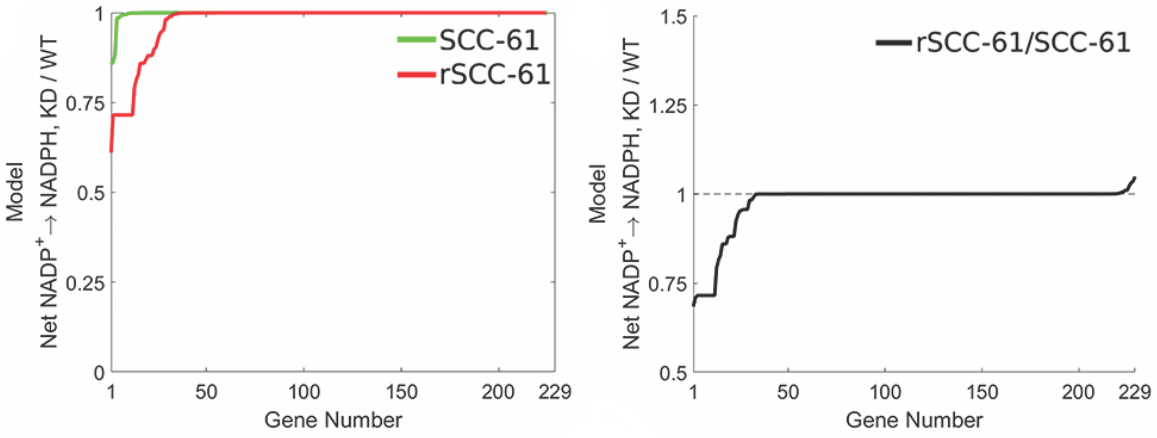


Figure 3-16. Model-predicted effect of gene knockdown on net conversion of $NADP_+$ to NADPH, expressed as a fraction of net conversion before knockdown. (Left) SCC-61 and rSCC-61 results separately. (Right) The ratio of rSCC-61/SCC-61 (values <1 signify gene knockdowns that lower NADPH production more in rSCC-61 cells, values >1 signify gene knockdowns that lower NADPH production more in SCC-61 cells).

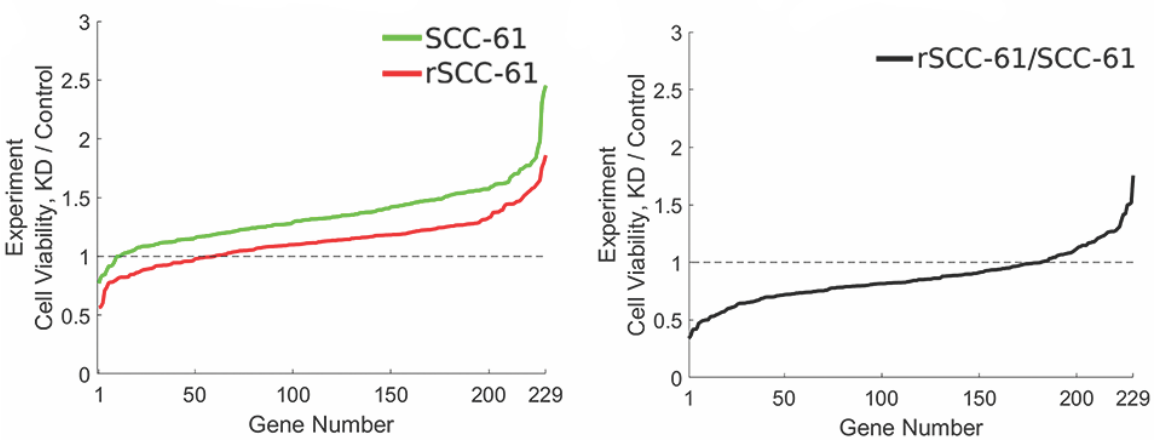


Figure 3-17. Experimental effect of gene knockdown on cell viability after β-lapachone treatment; the average value of three replicate experiments was taken, and the result expressed as a fraction of *Drosophila* control siRNA-transfected

cells. (Left) SCC-61 and rSCC-61 results separately; the *dotted line* represents division between genes that decrease *versus* increase cell viability after knockdown. (Right) The ratio of rSCC-61/SCC-61 (values <1 signify gene knockdowns that lower cell viability more, or raise cell viability less, in rSCC-61 cells; values >1 signify gene knockdowns that lower cell viability more, or raise cell viability less, in SCC-61 cells).

Six of the eighteen genes from Figure 2-4 were included in the siRNA panel; among these six genes, five displayed the same rSCC-61/SCC-61 relationship (>1 or <1) when comparing model NADPH production and experimental cell viability (Table 3-5). Aldehyde dehydrogenase ALDH1L1 and ALDH1L2, which had the third and second largest rSCC/SCC cell viability ratio among all 229 genes, also had the eighth and 12th largest rSCC/SCC NADPH production ratio. This result demonstrates the model's capability to predict β-lapachone sensitivity from cellular NADPH production for canonical NADPH-producing genes.

Table 3-5. Comparison of model and experimental results for canonical NADPH-producing genes (*left*) and discovered HNSCC-specific genes of interest (*right*).

Canonical NADPH-Producing RNAi Targets			HNSCC-Specific RNAi Targets of Interest		
Gene	Model rSCC-61/SCC-61	Experiment rSCC-61/SCC-61	Gene	Model rSCC-61/SCC-61	Experiment rSCC-61/SCC-61
G6PD	< 1	< 1	CYB5R1	1	< 1
MTHFR	< 1	< 1	CBR1	< 1	< 1
ME1	> 1	< 1	CYP4F3	< 1	< 1
IDH1	< 1	< 1	CAT	< 1	< 1
ALDH1L1	> 1	> 1	GSR	< 1	< 1
ALDH1L2	> 1	> 1	ALDH3A1	> 1	> 1

In addition, there was strong agreement between model and experimental results for genes not identified as major contributors to NADPH production, but nonetheless implicitly involved in metabolic differences between the two cell lines. The emergence of CYP4F3 as a gene affecting both model NADPH production and experimental β-lapachone sensitivity may stem from the tissue-specific expression of this monooxygenase in the trachea and esophagus as well as its reliance on NADPH as a cofactor for arachidonic acid utilization [184]. The strong effect of NADH-cytochrome b5 reductase (CYB5R1) levels on 24-h drug response (smallest experimentally determined rSCC-61/SCC-61 cell viability ratio) further confirms that rerouting of NADPH flux to enhance fatty acid synthesis serves as a critical metabolic difference between the cell lines as reported by Mims *et al.* [179], with knockdown in rSCC-61 cells increasing drug sensitivity (0.60 ± 0.28 cell viability ratio compared with control siRNA) while knockdown in SCC-61 cells protecting against chemotherapy (1.77 ± 0.42 cell viability ratio compared with control siRNA).

Surprisingly, silencing of CAT showed this differential trend as well to a lesser degree. CAT is typically associated with low expression in tumor cells that are β-lapachone

sensitive [160, 169], and the ratio of NQO1/CAT is critical for tumor-specific selectivity; thus, we expected that reduction of CAT would enhance redox cycling and reduce cell viability in both cell lines. Slightly reduced cell viability was observed in the CAT-silenced rSCC-61 cells along with slightly enhanced viability in the CAT-silenced SCC-61 cells, resulting in a pronounced rSCC-61/SCC-61 ratio of 0.53 and placing this perturbation in the top 10% of differentially affected genes. To explore this observation further, we evaluated the flux distributions of CAT and SOD1 to determine if concomitant changes in superoxide disproportionation could be the source of cell line-specific responses. We found no correlation in either cell line model between SOD1 and CAT flux values (SCC-61: $R_2 = 0.09$, rSCC-61: $R_2 = 0.02$); however, we did note that SOD1 flux values in the rSCC-61 model were constrained within a narrow range at a much higher flux than in SCC-61 cells (0.73 ± 0.08 mmol/gDW/h *vs.* 0.13 ± 0.11 mmol/gDW/h). Thus, rSCC-61 cells may be selected to a regime in which there is little to no margin for adjusting H_2O_2 production when H_2O_2 clearance becomes limited. GSR is another redox enzyme predicted to yield differential β-lapachone sensitivity by cell line-specific models and validated by experimental results. GSR provides a key couple between NAPDH/$NADP_+$ and GSH/GSSG in the cytosol to sustain ROS clearance through the glutathione/glutaredoxin antioxidant pathway. Glutathione S-transferase conjugates GSH adducts to β-lapachone hydroquinone for drug efflux from cells [185]; our results are consistent with GSR modulation limiting reduced glutathione for this process and thus sustaining intracellular drug levels for enhanced redox cycling.

Carbonyl reductase (CBR1), highly expressed in the oral mucosa [186], also demonstrated pronounced differences between the cell lines. Because CBR1 is a two-

electron reductase capable of reducing quinones [187], the increased sensitivity of rSCC-61 cells to β-lapachone with reduced CBR1 levels suggests a competing role of CBR1 in the β-lapachone redox cycling mechanism. Finally, not all perturbations resulted in more β-lapachone sensitivity and greater reduction of viability in the rSCC-61 cells. For example, ALDH3A1 is implicated in the enzymatic detoxification of aldehydes specifically in the gastrointestinal tract. Consistent with other ALDHs examined in this study (ALDH1L1 and ALDH1L2), silencing of this $NAD(P)^+$-consuming enzyme resulted in an rSCC-61/SCC-61 ratio >1.

3.4 Discussion

Because of NADPH involvement throughout the human metabolic network, genome-scale metabolic modeling is an ideal tool to investigate NADPH production between radiation-sensitive and radiation-resistant HNSCC phenotypes. Genome-wide metabolic models for NCI-60 cancer cell lines are being rapidly developed [188-190]; however, models for head and neck cancer cell lines that are not represented in the NCI-60 list are lacking. To establish HNSCC-specific models, we have developed a human genome-scale metabolic modeling platform that incorporates quantitative transcriptomic, kinetic, thermodynamic, and metabolite concentration data to create cell-type-specific models of radiation-sensitive (SCC-61) and radiation-resistant (rSCC-61) HNSCC cell lines (Figure 3-2). While model predictions are in close proximity to experimental measurements when both concentration and thermodynamic data are used, unsatisfactory predictions result when one or both data sources are not included, demonstrating the necessity of integrating

disparate biological data types into the metabolic model to achieve valid results (Figure 3-5).

A common practice used in FBA models is treating gene expression data as binary (reaction on or off) to achieve cell-type-specific results. Proteomic data better represent gradations in enzyme levels than gene expression data, and setting reaction upper bounds relative to quantitative proteomic values is more accurate biologically than completely removing reactions from the model that fall under some arbitrary threshold. For our model construction, we leveraged the transcriptional similarities and differences between the matched SCC-61/rSCC-61 cell lines and the CAL-27 head and neck cancer cell line to estimate proteomic values for SCC-61 and rSCC-61 cells. This approach maintains known proteomic–transcriptomic correlations where this information was available and provides accurate predictions of NADPH production.

β-Lapachone is a chemotherapeutic drug that has tremendous promise in increasing the radiation sensitivity of cancer cells through an NADPH-consuming futile redox cycle and depletion of cellular NADPH stores. Since antioxidative and DNA repair mechanisms in radiation-resistant tumors rely on sufficient NADPH supply, combining radiation therapy with β-lapachone treatment may be an effective strategy for cancer patients who previously have not responded to radiation. Because of the interconnection between β-lapachone's mechanism of action and cellular NADPH levels, our hypothesis was that the relative β-lapachone sensitivity tested with the siRNA screen of oxidoreductase genes on radiation-susceptible and radiation-resistant cells could be predicted by the impact of gene knockdown on total cellular NADPH. To test this hypothesis, we combined our novel genome-scale metabolic modeling platform with the cellular objective function of

maximizing conversion from $NADP^+$ to NADPH while maintaining sufficient ATP production to obtain biologically viable results; this differs from the classical cancer objective functions of maximizing biomass or ATP production. We justify our choice of a novel objective function by noting the generation of the rSCC-61 cell line: SCC-61 cells were subjected *in vitro* to repeated rounds of ionizing radiation (2 Gy for a cumulative total of 16 Gy) to select radiation-resistant cells; thus, the cells optimized to repair double-stranded DNA breaks through the use of NADPH were more likely to survive. The interconnectedness of NADH, NAD^+, NADPH, and $NADP^+$ in maintenance of the global nicotinamide pool warrants future exploration in the use of alternative objective functions, such as optimization of NAD^+ in the radioresistant phenotype. For example, in all $NQO1^+$cancer cells examined to date (pancreatic, breast, prostate, NSCLC, and others), NAD^+ loss correlates well with loss of survival monitored by short-term or long-term assays. Because the two cell lines are so extensively characterized for epigenetic, transcriptomic, proteomic, and metabolomic changes occurring with gain of a radiation resistance phenotype, they serve as a useful system for testing the underlying hypothesis that cumulative alterations across the genome that yield changes in NADPH production to facilitate radiation-induced DNA damage repair will also enhance NADPH-dependent redox cycling for increased β-lapachone sensitivity. Synergy between ionizing radiation therapy and β-lapachone treatment has been established in head and neck cancer [167]; our flux balance models suggest specific mechanisms that result in this effect.

We first found through our modeling that different NADPH-producing reactions contribute to the total cellular NADPH pool on varying flux scales (Figure 3-6). Unlike a prior FBA model that strongly implicated the 10-formyl-tetrahydrofolate pathway and

malic enzyme in NADPH production for HEK293 cells [144], reactions in the α-ketoglutarate production and pentose phosphate pathways appear to contribute more to the NADPH pool for head and neck cancer cells. In addition, whereas most reaction flux distributions did not differ much between SCC-61 and rSCC-61 cells, flux values through the GLUD1/GLUD2 reaction were greatly increased in rSCC-61 cells compared with SCC-61. Even though some reactions produce more NADPH than others, the amount of NADPH production did not appear to correlate with the effect of simulated gene knockdown on total cellular NADPH (Figure 3-7). For example, knockdown of GLUD1 or GLUD2 in the models had very little effect on total NADPH production in SCC-61 and rSCC-61 cells. In contrast, knockdown of IDH1, G6PD, and ME2 differentially impacted NADPH production in rSCC-61 cells, and knockdown of IDH2 differentially impacted SCC-61 cells. Collectively, these results suggest that the metabolic network can reroute NADPH-producing flux to other pathways in a cell-type-specific manner to achieve the goal of maximizing NADPH production. This was verified with the differential effect of IDH1 knockdown on rerouting flux through GLUD1/GLUD2 between SCC-61 and rSCC-61 cells (Figure 3-8).

Although our modeling approach represents a vast improvement over past human genome-scale metabolic models, it still lacks some detail regarding redox-based reactions. Given the dramatic changes in flux distributions upon simulated IDH silencing, inclusion of oncometabolite-generating reactions, such as recently reported for IDH mutations in HNSCC [191], would be informative for prediction of quinone drug redox cycling. Another limiting factor in genome-scale metabolic modeling is the lack of proteomic, kinetic, thermodynamic, and concentration parameters for all reactions and metabolites within

human metabolism. The FBA modeling platform is constraint based, meaning that more accurate predictions are obtained through the incorporation of more constraints on cellular function. With greater availability of biological parameters to include within genome-scale models, more constraints can be imposed on cellular function to improve the predictive power of genome-scale models. This necessitates collaborative efforts to document experimentally verified parameter values on a genome scale. Furthermore, a limitation of FBA modeling is the steady-state assumption required for calculation of metabolic flux values. Many of the oxidoreductase enzymes under discussion in this study are regulated by the redox-sensing transcription factor, Nrf2. In particular, NQO1 contains an antioxidant response element sequence in the promoter region and, due to its high induction, is often used as a positive control for Nrf2-regulated gene activation [192]. A feature of β-lapachone cytotoxicity is that the generation of intracellular ROS by futile quinone cycling can trigger Nrf2, thereby increasing NQO1 levels and augmenting the initial drug mechanism of action. We did not observe changes in NQO1 levels in either cell line during a period of 8 h post-β-lapachone exposure; however, other enzyme expression changes during this time frame may alter metabolic flux values from our baseline analysis. Nonetheless, the snapshot of genome-wide metabolism determined by our systems approach highlights the importance of basal NADPH flux distributions in the predictive value for NQO1 bioactivated chemotherapeutic compounds.

This systems biology analysis on a matched model of HNSCC radiation response provides insight on how small variations in cell line transcriptomics and proteomics can cumulatively yield dramatic changes in metabolism and large phenotypic changes to a therapeutic treatment. Our modeling efforts demonstrate that depending on the distribution

of flux, perturbations in metabolic enzymes can have opposite effects on the NADPH/$NADP^+$ redox couple, the degree of futile redox cycling of a chemotherapeutic, and ultimately the sensitivity of cell viability to the drug. This modeling approach allows for continued investigations on synergy of redox-based chemotherapies and radiation therapy.

CHAPTER 4. FBA MODELS OF REDOX METABOLISM IN TCGA TUMORS

4.1 Introduction

To investigate the underlying pathophysiological mechanisms of radiation resistance and discover targets for improving sensitivity to radiation therapy, previous studies have compared differences in tumor biology between radiation-sensitive and radiation-resistant patients. Correlative studies using single omics modalities such as genomics or transcriptomics have facilitated the discovery of predictive biomarkers and gene expression signatures for radiation sensitivity [3, 46, 171, 193-195]. However, these previous findings fall short by failing to integrate multiple biological data types, analyze differential expression in the context of genome-scale biochemical and regulatory networks, or provide mechanistic insights into how these prospective biomarkers impact tumor function and can be exploited to improve radiation sensitivity.

Ionizing radiation therapy results in the generation of reactive oxygen species (ROS) such as superoxide (O_2^-) and hydrogen peroxide (H_2O_2), which oxidize the cellular environment and damage cellular structures including DNA [196-198]. A reduced intracellular environment with elevated levels of NADPH, NADH, and GSH is favored by tumor cells for resistance to ROS-mediated damage caused by ionizing radiation. For example, NADPH is used by glutathione reductase and thioredoxin reductase to reduce GSH and thioredoxin, respectively, which comprise two of the most important antioxidant systems [199]. Glutathione peroxidase and peroxiredoxins utilize GSH and reduced thioredoxin, respectively, to lower intracellular ROS levels [200]. Additionally, NADPH

and NADH are cofactors for dihydrofolate reductase and dihydroorotate dehydrogenase, respectively, which are enzymes involved in nucleotide production; thus, increased levels of these cofactors in turn increase the supply of deoxynucleotides for DNA damage repair in response to ionizing radiation [57, 201]. GSH is also directly involved in DNA damage repair by reduction of nucleotide bases [202, 203]. Because redox metabolism is critical to the response of tumors to ionizing radiation, identifying targets for inhibiting production of these redox cofactors may provide a valuable strategy for sensitizing tumors to radiation therapy.

Redox metabolites such as NADPH are ubiquitous throughout the human metabolic network. In the most recent reconstruction of the human metabolic network, it is estimated that 957 metabolic reactions directly involve $NADP^+$ and/or NADPH as cofactors [7]. Additionally, the fluxes through metabolic pathways connecting and feeding into these reactions dictate the observed activity of NADPH-consuming and -generating reactions, making it necessary to study the entire metabolic network to obtain an accurate understanding of redox metabolism. Because using experimental methods to study all 13,000+ metabolic reactions, their interconnections, and responses to perturbations is currently infeasible, computational methods of modeling human metabolism are needed. Flux balance analysis (FBA) is a computational approach for developing genome-scale human metabolic models and predicting steady state reaction fluxes for individual samples of interest [9]. By combining the stoichiometric representation of the human metabolic network, constraints on the fluxes through metabolic reactions, and an objective function to maximize a particular metabolic phenotype, predictions of maximum reaction fluxes or metabolite production rates under physiological constraints can be obtained. These

predictions can be compared between samples of interest to investigate metabolic differences contributing to pathophysiological processes.

Recently, FBA models personalized to individual cell types or tumors have been developed through the integration of transcriptomic data; however, there remain significant methodological shortcomings in the construction of these models, which have hindered their ability to yield accurate and quantitative metabolic predictions [11]. Commonly-used FBA algorithms such as GIMME, iMAT, MADE, and CORDA utilize gene expression information to determine which metabolic reactions are "on" (reactions are kept in the model) or "off" (reactions are removed from the model) based on the expression of the associated metabolic enzyme [148, 149, 152, 204]. However, many of these methods rely on arbitrary thresholding of gene expression values to determine on and off states, and the complete removal of reactions with low associated gene expression is not physiologically accurate. Other methods such as E-Flux constrain the maximum flux through a particular reaction in proportion to the associated gene expression, but arbitrary reference values are still used to set loose versus tight constraints on predicted fluxes [150]. Additionally, most FBA models fail to incorporate any kinetic or thermodynamic constraints, which greatly affect metabolic fluxes and the directionalities of individual reactions [12]. These shortcomings in model construction have ultimately limited the clinical utility of FBA models for accurately predicting metabolic phenotypes and directly improving the diagnosis and treatment of disease [205].

We have previously developed FBA models of radiation-sensitive and radiation-resistant head and neck squamous cell carcinoma cell lines through the integration of transcriptomic, kinetic, and thermodynamic information into quantitative constraints on the

maximum flux and directionality of metabolic reactions [206]. These models accurately predicted response to treatment with the NADPH-dependent redox-cycling chemotherapeutic β-lapachone and facilitated the discovery of novel oxidoreductase genes that differentially affected β-lapachone response between radiation-sensitive and -resistant cancer cells. Additionally, we identified that radiation-resistant cancer cells were capable of increasing fluxes through NAD salvage and purine salvage pathways to increase production of NADH; predicted NAD-generating fluxes were greatly increased through NAMPT, an enzyme whose activity has previously been associated with radiation resistance and poor survival in cancer patients [79, 207]. Here, we extend this approach by developing an automated bioinformatics pipeline for integrating multi-omic information into personalized FBA models of 716 radiation-sensitive and 199 radiation-resistant patient tumors across multiple cancer types from The Cancer Genome Atlas (TCGA) [2]. These models are used to investigate differences in redox metabolism between radiation-sensitive and -resistant tumors, identify patient clinical factors that are associated with particular metabolic phenotypes, and discover novel targets redox metabolism in radiation-resistant tumors.

4.2 Methods

4.2.1 Computational methods

4.2.1.1 Objective functions

Table 4-1 lists the objective functions used in FBA and FVA. To maximize the production of a metabolite in all cellular compartments, separate objective functions for each compartment with the same weight were simultaneously maximized. To ensure that

biologically viable solutions were obtained, all models were checked to ensure that they were capable of producing physiological ATP levels typical of mammalian cells (1.0625 mmol gDW^{-1} hr^{-1}) while maximizing the imposed objective functions [208].

Table 4-1. Objective functions used in FBA and FVA.

Title	Objective Function
$NADP^+$ → NADPH	1 nadph[all] → 1 nadp[all] + 1 h[all]
$NADP^+$ → NADPH, Cytosolic	1 nadph[c] → 1 nadp[c] + 1 h[c]
$NADP^+$ → NADPH, Mitochondrial	1 nadph[m] → 1 nadp[m] + 1 h[m]
NAD^+ → NADH	1 nadh[all] → 1 nad[all] + 1 h[all]
GSSG → GSH	2 gthrd[all] + 1 nadp[all] → 1 gthox[all] + 1 nadph[all] + 1 h[all]
H_2O_2 →	Ø → 1 h2o2[all]
→ dNTP	1 datp[all] + 1 dctp[all] + 1 dgtp[all] + 1 dttp[all] → Ø
→ Palmitate	1 hdca[all] → Ø

Note: "[all]" signifies that the objective function was maximized in all cellular compartments with equal weights. "Ø" signifies no metabolites on the left or right side of the equation.

4.2.1.2 Metabolic network

Recon 3D version 3.01 was used as the core metabolic network [7]. While the "full" version (Recon3D_301.mat) produced erroneous energy-generating cycles (e.g. producing ATP without uptake of any metabolic precursors), the "model" version (Recon3DModel_301.mat) was found not to be the largest subset of reactions without

erroneous energy-generating cycles (additional reactions from the "full" version could be added). To this end, following the procedure outlined in Fritzemeier et al., reactions from the "full" version were sequentially added back into the "model" version as long as they did not result in erroneous energy-generating cycles [209].

To address missing and inaccurate redox-based reaction information within Recon3D, the following changes were made:

1. The reaction "FTHFDH" was split into two separate cytosolic and mitochondrial reactions. The cytosolic reaction is catalyzed by ALDH1L1, and the mitochondrial reaction is catalyzed by ALDH1L2.

   Originally:

   a. 1 10fthf[c] +1 h2o[c] + 1 nadp[c] → 1 co2[c] + 1 h[c] + 1 nadph[c] + 1 thf[c], GPR: 10840.1 or 160428.1, EC: 1.5.1.6

   Updated:

   a. 1 10fthf[c] + 1 h2o[c] + 1 nadp[c] → 1 co2[c] + 1 h[c] + 1 nadph[c] + 1 thf[c], GPR: 10840.1, EC: 1.5.1.6
   b. 1 10fthf[m] + 1 h2o[m] + 1 nadp[m] → 1 co2[m] + 1 h[m] + 1 nadph[m] + 1 thf[m], GPR: 160428.1, EC: 1.5.1.6

2. MTHFR was removed from the GPR associated with reaction "MTHFD". MTHFR catalyzes the reaction converting 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which is already included in Recon3D.

   Originally:

   a. 1 mlthf[c] + 1 nadp[c] ←→ 1 methf[c] + 1 nadph[c], GPR: 4522.1 or 4524.1, EC: 1.5.1.5

Updated:

a. 1 mlthf[c] + 1 nadp[c] ←→ 1 methf[c] + 1 nadph[c], GPR: 4522.1, EC: 1.5.1.5

3. Reactions catalyzed by isoforms of NADPH oxidase were added.

   Originally: None

   Updated:

   a. 1 nadph[c] + 2 o2[c] → 2 o2s[e] + 1 nadp[c] + 1 h[c], GPR: 27035.1 or 1536.1 or 50508.1 or 79400.1, EC: 1.6.3.1
   b. 1 nadph[n] + 2 o2[n] → 2 o2s[c] + 1 nadp[n] + 1 h[n], GPR: 50507.1, EC: 1.6.3.1
   c. 1 nadph[r] + 2 o2[r] → 2 o2s[c] + 1 nadp[r] + 1 h[r], GPR: 50507.1, EC: 1.6.3.1

4. PRDX1 and PRDX2 were removed from the GPR associated with reaction "GTHP". Glutathione peroxidase involves the reduction of hydrogen peroxide by glutathione, which does not involve peroxiredoxins. Peroxiredoxin reactions are added to Recon3D (see below)

   Originally:

   a. 1 h2o2[c] + 2 gthrd[c] → 2 h2o[c] + 1 gthox[c], GPR: 7001.3 or 5052.3 or 2877.1 or 2876.2 or 5052.2 or 2876.1 or 5052.1 or 7001.1 or 2879.1 or 7001.2, EC: 1.11.1.9

   Updated:

   a. 1 h2o2[c] + 2 gthrd[c] → 2 h2o[c] + 1 gthox[c], GPR: 2877.1 or 2876.2 or 2876.1 or 2879.1, EC: 1.11.1.9

5. Similarly to (4), PRDX3 was removed from the GPR associated with reaction "GTHPm".

   Originally:

   a. 1 h2o2[m] + 2 gthrd[m] → 2 h2o[m] + 1 gthox[m], GPR: 10935.1 or 10935.2 or 2879.1 or 2876.1, EC: 1.11.1.9

   Updated:

   a. 1 h2o2[m] + 2 gthrd[m] → 2 h2o[m] + 1 gthox[m], GPR: 2879.1 or 2876.1, EC: 1.11.1.9

6. PRDX6 was removed from the GPR's associated with the following reactions: HMR_0960, HMR_0963, HMR_0988, HMR_2441, HMR_1048, HMR_1049. These are all glutathione peroxidase reactions, which PRDX6 is not involved in.

7. Reactions catalyzed by isoforms of peroxiredoxin were added:

   Originally: None

   Updated:

   a. 1 h2o2[c] + 1 trdrd[c] → 2 h2o[c] + 1 trdox[c], GPR: 5052.1 or 5052.2 or 5052.3 or 7001.1 or 7001.2 or 7001.3, EC: 1.11.1.15

   b. 1 h2o2[m] + 1 trdrd[m] → 2 h2o[m] + 1 trdox[m], GPR: 10935.1 or 10935.2, EC: 1.11.1.15

8. The oxidation and glutathionylation of protein thiol groups, as well as their reduction by glutaredoxin and thioredoxin, were added:

   Originally: None

   Updated:

a. 1 h2o2[c] + 1 Pr-SH[c] + 1 gthrd[c] → 2 h2o[c] + 1 Pr-SSG[c], GPR: None, EC: None
b. 1 h2o2[m] + 1 Pr-SH[m] + 1 gthrd[m] → 2 h2o[m] + 1 Pr-SSG[m], GPR: None, EC: None
c. 1 Pr-SSG[c] + 1 gthrd[c] → 1 Pr-SH[c] + 1 gthox[c], GPR: 2745.1, EC: 1.20.4.1
d. 1 Pr-SSG[m] + 1 gthrd[m] → 1 Pr-SH[m] + 1 gthox[m], GPR: 2745.1, EC: 1.20.4.1
e. 1 h2o2[c] + 1 Pr-SH[c] → 2 h2o[c] + 1 Pr-SS[c], GPR: None, EC: None
f. 1 h2o2[m] + 1 Pr-SH[m] → 2 h2o[m] + 1 Pr-SS[m], GPR: None, EC: None
g. 1 Pr-SS[c] + 1 trdrd[c] → 1 Pr-SH[c] + 1 trdox[c], GPR: None, EC: None
h. 1 Pr-SS[m] + 1 trdrd[m] → 1 Pr-SH[m] + 1 trdox[m], GPR: None, EC: None

9. The reliance of catalase enzyme on NADPH to prevent oxidation of its active site by H_2O_2 was incorporated [69]. Kirkman et al. measured rates of H_2O_2 formation and corresponding rates of NADPH oxidation in mixtures of catalase and varying concentrations of glucose oxidase [210]. If the rate of H_2O_2 formation by glucose oxidase is assumed to equal the rate of H_2O_2 clearance by catalase, then the slope of NADPH oxidation vs. H_2O_2 formation gives the amount of NADPH oxidized by catalase per molecule of H_2O_2 cleared. Figure 4-1 shows that for every 1 molecule of H_2O_2 cleared by catalase, 0.0641 molecules of NADPH are oxidized; thus, the chemical equation for catalase activity is:

$$2\ H_2O_2 + 0.1282\ \mathrm{NADPH} \rightarrow 2\ H_2O + 1\ O_2 + 0.1282\ \mathrm{NADP^+} + 0.1282\ H^+ \quad (23)$$

This change was made for catalase reactions in the cytosol, mitochondria, peroxisome, and endoplasmic reticulum.

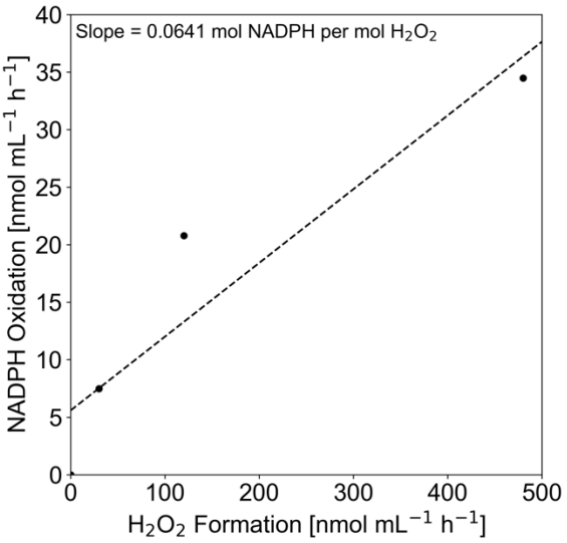


Figure 4-1. Regression of NADPH oxidation by catalase enzyme vs. H_2O_2 formation by varying concentrations of glucose oxidase. Data from [210].

10. The neomorphic 2-hydroxyglutarate-producing reaction catalyzed by mutant IDH1 was added. Reactions for 2-hydroxyglutarate export were also added.

    Originally: None

    Updated:

    a. 1 akg[c] + 1 nadph[c] → 1 M00653[c] + 1 nadp[c], GPR: 3417.1, EC: 1.1.1.42
    b. 1 M00653[c] → 1 M00653[e], GPR: None, EC: None
    c. 1 M00653[e] → , GPR: None, EC: None

Because transcriptomic and proteomic data were available at the individual gene level but not at the individual isoform level, all isoforms for each individual gene within Recon3D were combined.

Other changes made to Recon3D include:

- Removing the following duplicate reactions (where another reaction exists where the stoichiometric vector is an exact multiple of these duplicate reactions):
  "HMR_7257", "PEHSFABPe", "G6PDH2c", "GNDc", "PGLc", "RPEc"
- Removing EC numbers from all transport reactions, since many of these were incorrect.
- Removing the following genes and any associations with them within Recon3D reactions, since information about these genes could not be found:
  NCBI Gene ID's: 0, 100507855, 8041
- Removing reactions within the following subsystems, as this was necessary to prevent de novo production of essential amino acids:
  "Protein assembly", "Protein degradation", "Protein modification", "Protein formation"
- Removing reactions in the "Bile acid synthesis" subsystem, as these are not expected to be active in most tumors
- Removing NQO1 futile cycle reactions, as these un-constrained reactions resulted in erroneous predictions of NADPH production: "NADPQNOXR", "NADQNOXR", "HMR_9538", "HMR_9722"

- Removing oxalosuccinic acid-producing IDH1 reactions, as these overlapped with α-ketoglutarate-producing IDH1 reactions: “r0423”, “r0424”, “r0425”, “r0422”

4.2.1.3 Model flux constraints

Flux bounds for reactions within Recon3D with both associated GPR rules and EC numbers were constrained using the following Michaelis-Menten upper flux bound equation:

$$v_{max} = k_{cat}[E] \tag{24}$$

where k_{cat} and [E] are the turnover number (units of hr^{-1}) and abundance (units of mmol gDW^{-1}) of the associated enzyme catalyzing that reaction, respectively. By using the abundances of enzymes within individual tumors, these models were able to make metabolic predictions that are specific to the individual sample of interest. Additionally, these Michaelis-Menten constraints allow models to make flux predictions that are quantitative and physiologically-accurate.

Equation 24 was used to set upper flux bounds $v_{max,j}$ for all reactions with associated GPR rules and EC numbers, as well as lower flux bounds $v_{min,j}$ for such reactions that are reversible (where $v_{min,j} < 0$). See sections “Enzyme Abundance Calculation” and “Turnover Number Calculation” for more information.

4.2.1.4 Enzyme abundance calculation

RNA-Seq gene expression data from TCGA patient tumors was obtained from Rahman et al.’s alternative preprocessing method (Gene Expression Omnibus GSE62944) [211]. Data from this preprocessing method showed fewer missing values, more consistent expression between replicates, and improved prediction of biological pathway activity compared to the original TCGA pipeline.

Enzyme abundances within individual samples were predicted from sample gene expression data using the pipeline in Figure 4-2. Although this approach does not perfectly predict experimentally-measured protein abundances or activities, it allows for the generation of FBA predictions that are more quantitative and physiologically-accurate than other methods that rely solely on thresholding of gene expression data. The approach described below relies only on the gene expression data from the individual sample of interest and not data from other samples included within the dataset, improving the reproducibility of this method as well as allowing for new samples to be analyzed post-hoc.

Figure 4-2. Pipeline for the estimation of enzyme abundance values from gene expression data for individual TCGA patient tumors. Percentages represent the percentage of enzyme abundances of associated genes in Recon3D that have been cumulatively determined at each step in the pipeline for all TCGA samples.

The individual steps in the pipeline are explained herein:

1. *Schwanhäusser et al.* - By measuring the transcription, translation, and degradation rates of individual genes and proteins in NIH3T3 cells, Schwanhäusser et al. developed an ODE model relating mRNA and protein abundances per cell that encompasses 955/3268 (29.2%) genes in Recon3D [212]. Their model also took into account proteins with half-lives longer than the length of the cell cycle, and how this would impact measurements of protein abundance. The authors demonstrated that these rate constants and model could be applied to accurately

predict protein abundances in other samples including cancer cell lines. Using this model, gene expression values (transcripts per kilobase-million (TPM)-normalized) were converted to protein abundance values (parts per million (PPM)-normalized) for the 955 Recon3D genes included.

2. *Linear regression* - For each sample, a linear regression of predicted protein abundance using the Schwanhäusser et al. method versus the measured gene expression values of corresponding genes is performed (example is shown in Figure 4-3). This regression model is used to predict the abundance of proteins where an associated gene expression value is available but not corresponding parameter values in the Schwanhäusser et al. method.

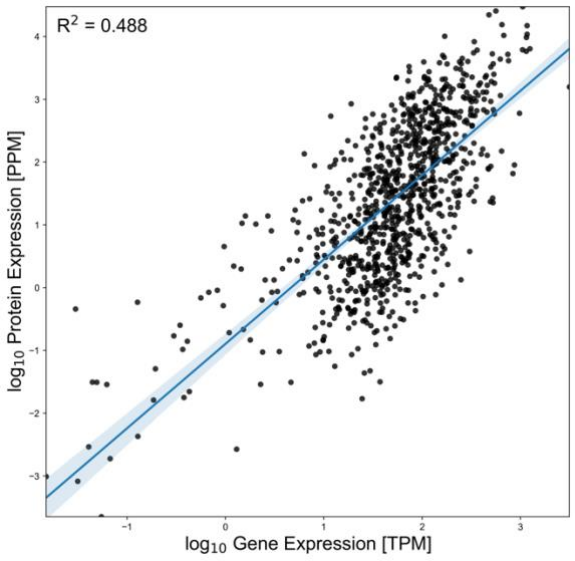


Figure 4-3. Example of linear regression step in the pipeline for enzyme abundance values, where predicted protein abundances using the Schwanhäusser et al. method are regressed versus the measured gene expression values of corresponding genes in an individual TCGA sample.

3. *PaxDB* - If no gene expression value is available for a particular sample, the harmonic mean of abundance values for the associated protein across all *Homo sapiens* samples within the PaxDB database is taken [213]. If the associated protein is not available in the PaxDB database, the harmonic mean of abundance values for all proteins across all *Homo sapiens* samples is taken. Imposing the same values for all samples with missing gene expression values acts to prevent differences in FBA model predictions between samples due to missing data. The harmonic mean of PaxDB values is taken since (1) the arithmetic mean tends to overestimate representative values of positively-skewed distributions (as seen with protein abundance values); and (2) the smaller protein abundance estimate of the harmonic mean well characterizes the fact that missing measured RNA-Seq values are usually at least in part due to low actual mRNA expression.

For genes with associated enzymes in multiple cellular compartments (e.g. catalase, which localizes to the cytosol, mitochondria, peroxisome, and endoplasmic reticulum), the predicted protein expression value was divided into each cellular compartment with weights proportional to the COMPARTMENTS value for the cellular compartment in which the metabolic reaction takes place [214]. These COMPARTMENTS values provide a score representing the confidence that a particular protein is found within a particular cellular compartment, using both experimental and computational evidence.

Validation of this pipeline was performed to ensure (1) that the predicted enzyme abundance values correlated at least as well with experimental protein expression data from TCGA than gene expression data; and (2) that the predicted enzyme abundance values were

at physiological orders of magnitude. Normalized RPPA experimental protein expression values from TCGA samples were available for 30 genes/proteins in Recon3D [2]. The correlation between experimental enzyme abundance and predicted enzyme abundance values ($R^2_{PredictedAbundance}$) was greater than the correlation between experimental enzyme abundance and gene expression values (R^2_{Gene}) for 18/30 genes (60%; Figure 4-4, Figure 4-5). Additionally, the largest change in correlation for genes where $R^2_{PredictedAbundance} < R^2_{Gene}$ was less than 0.004, indicating that the $R^2_{PredictedAbundance}$ was not significantly lower than the R^2_{Gene} for any of the 30 genes.

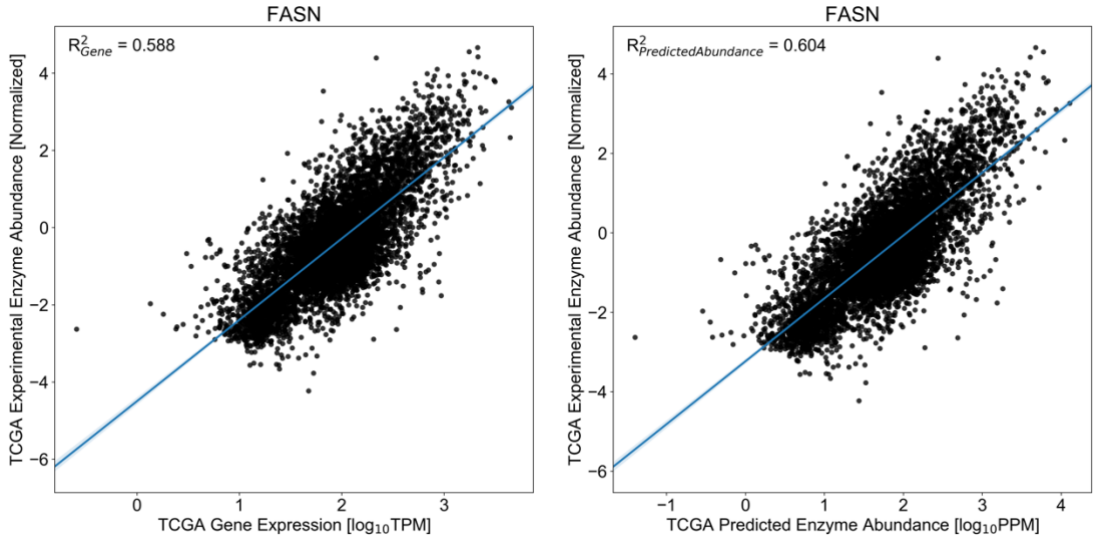


Figure 4-4. Example comparison of (left) gene expression and (right) predicted enzyme abundance values to experimental enzyme abundances from TCGA samples for an individual gene/protein (FASN). Blue line: least squares linear regression line.

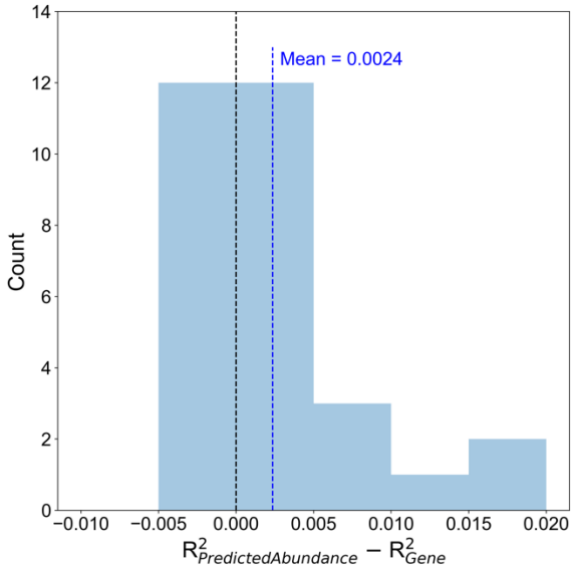


Figure 4-5. Improvement of correlation to experimental enzyme abundances of predicted abundance values ($R^2_{PredictedAbundance}$) compared to original gene expression data (R^2_{gene}) for all 30 genes/proteins with available experimental values for TCGA samples. Blue dashed line: mean value of $R^2_{PredictedAbundance}$-R^2_{gene} across all 30 genes/proteins.

Because experimental non-normalized protein abundance values were not available for TCGA samples for validation, experimentally-measured protein abundances in the NCI-60 panel of cancer cell lines were used to analyze the improvement in correlation to experimental data and physiological accuracy of predicted abundances on a genome scale (Figure 4-6, Figure 4-7). The mean values of predicted enzyme abundances across TCGA samples showed significantly greater correlation to experimental NCI-60 abundance values (R^2 = 0.450) compared to TCGA gene expression values (R^2 = 0.314). Additionally, the order of magnitude of predicted abundance values matched well with experimental values for individual genes/proteins over the majority of orders of magnitude. The match to

experimental values was greatest for enzyme abundances that were calculated at the first step of the pipeline (Schwanhäusser et al. method; R_2 = 0.586) compared to abundances calculated at later steps (R_2 = 0.286), indicating the necessity for future transcription, translation, and degradation rate measurements to be conducted for genes/proteins not covered in the Schwanhäusser et al. method.

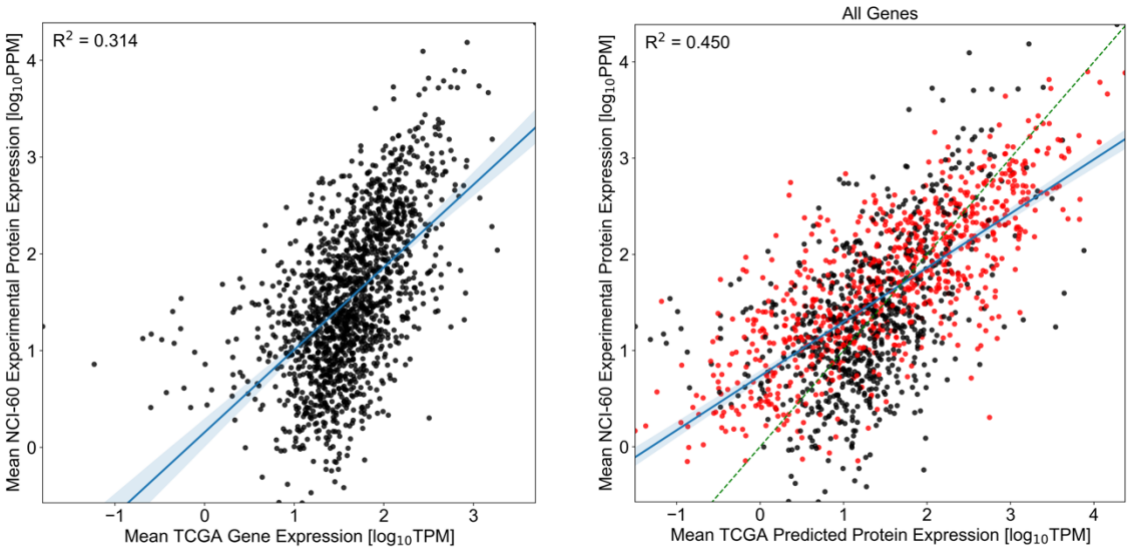


Figure 4-6. (Left) Correlation between mean gene expression values across TCGA samples and mean experimental enzyme abundances across NCI-60 samples for all 3,268 genes/proteins in Recon3D. Blue line: least squares linear regression line. (Right) Correlation between mean predicted enzyme abundances across TCGA samples and mean experimental enzyme abundances across NCI-60 samples for all 3,268 genes/proteins in Recon3D.

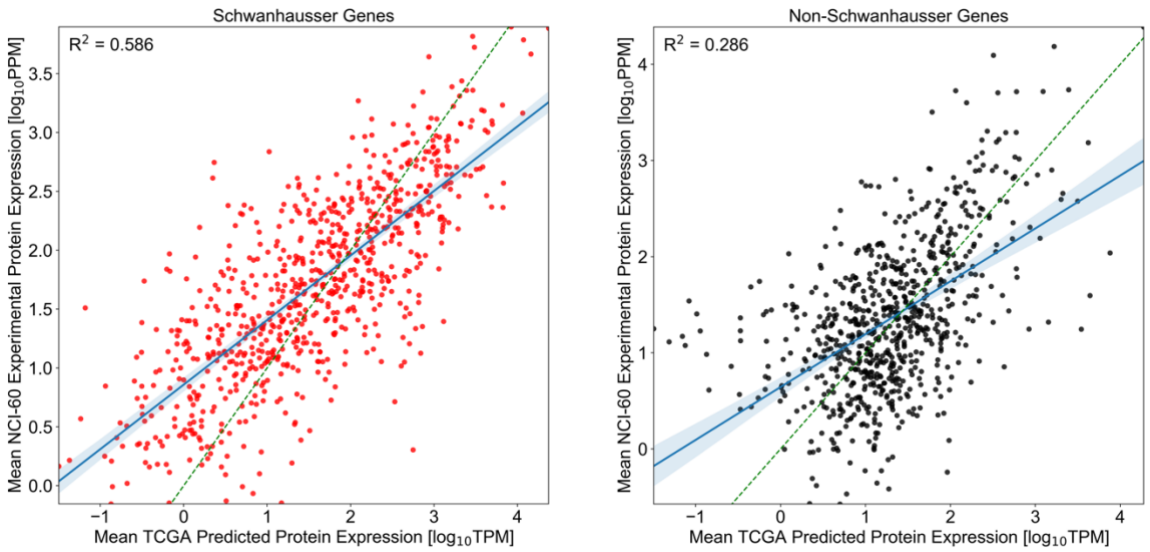


Figure 4-7. Breakdown of predicted enzyme abundances into (Left; red dots) those values calculated during the first step of the prediction pipeline (Schwanhäusser et al. method) and (Right; black dots) those values calculated at later steps of the pipeline. Green dashed line: 1:1 line.

4.2.1.5 Turnover number calculation

Enzyme turnover numbers for associated reactions within Recon3D were determined using available experimental data from the BRENDA database API [137]. Because the BRENDA database does not contain data for all Recon3D enzymes with the exact same enzyme commission (EC) number and substrate measured using human enzymes at 37°C and the correct cellular compartment-specific pH, a pipeline was created where Recon3D turnover values are filled in at subsequent steps with less accurate but more available information within BRENDA (Figure 4-8). For an individual enzyme-catalyzed reaction within Recon3D, the associated turnover number was determined as close to the start of the pipeline as data was available within BRENDA. Values determined towards the start of the pipeline are more accurate, but less turnover number data with this

greater accuracy is available within BRENDA. In determining turnover number accuracy, priority was given in the order of (1) using the correct EC number (versus using data from EC numbers that match only the first 3, 2, or 1 digits with that of the Recon3D reaction); (2) using the correct substrate as given in the Recon3D reaction (versus using data from any available substrate); (3) using turnover number data taken from human enzymes (versus using data from any available organism); and (4) using turnover number data taken at physiological temperature and pH (versus using data near but not at physiological temperature and pH). At a particular step in the pipeline, if multiple turnover number values that met the pipeline criteria were available, a weighted average of these values was taken, where the weights are equal to the Tanimoto coefficient (a measure of molecular similarity; greater coefficient signifies greater similarity between two molecules) between the desired substrate in the Recon3D reaction and the substrate in BRENDA for which the associated turnover number was determined. No turnover number values in the BRENDA database from mutated enzymes were used at any step in the pipeline.

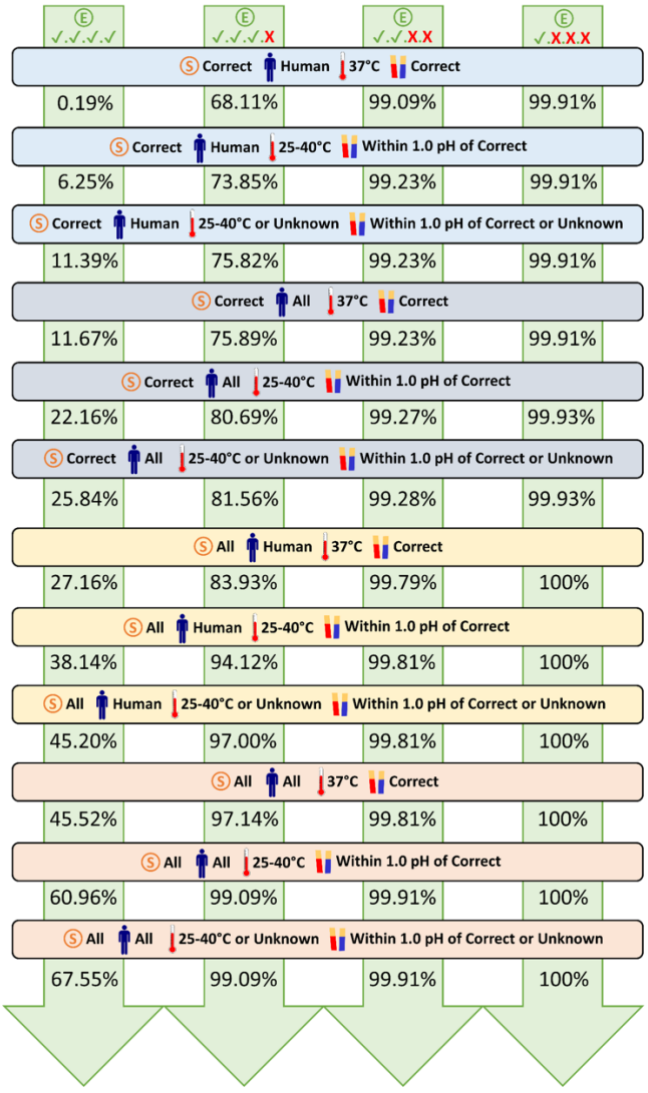


Figure 4-8. Pipeline for the estimation of enzyme turnover numbers using data from the BRENDA database API. Percentages represent the percentage of turnover numbers of associated reactions in Recon3D that have been cumulatively determined at each step in the pipeline.

4.2.1.6 Mutation data

The effect of single nucleotide polymorphisms (SNP's) located in Recon3D genes on the catalytic activity of the associated metabolic enzyme was estimated using the

Envision computational platform developed by Gray et al [137]. Envision leverages large-scale experimental mutagenesis datasets to predict a score representing the effect of single amino acid changes on protein function; a score less than 1 signifies a loss-of-function, while a score greater than 1 signifies a gain-of-function. These quantitative scores were shown to correlate better with experimentally-measured percent changes in protein activity after mutation than previously-developed methods including SNAP2 and EVmutation. For each TCGA patient with mutation data available, the turnover number for a particular Recon3D reaction was multiplied by the Envision score for any SNP's located in the associated gene.

4.2.1.7 IDH1 mutations

Particular mutations in isocitrate dehydrogenase 1 (IDH1), predominantly at amino acid R132, result in both the loss of function of its normal NADPH-generating activity, as well as the gain of function of a neomorphic NADPH-consuming activity that produces the oncometabolite 2-hydroxyglutarate [215]. This change in function was implemented by adding the IDH1-catalyzed neomorphic reaction to Recon3D and imposing turnover numbers of the normal and neomorphic reactions measured by Avellaneda Matteo et al. based on the IDH1 mutation status of each patient (Table 4-2) [216].

Table 4-2. Turnover numbers of normal and neomorphic IDH1 reactions for given IDH1 mutations.

Mutation	k_{cat} Normal [s^{-1}]	k_{cat} Neomorphic [s^{-1}]
WT	85 ± 4	0.019 ± 0.001
R132C	4.4 ± 0.1	1.60 ± 0.07
R132G	9.3 ± 0.6	1.59 ± 0.09
R132H	2.4 ± 0.1	4.2 ± 0.3

4.2.1.8 Thermodynamic constraints

Only metabolic reactions with a negative change in Gibbs free energy (ΔG) can carry a non-zero net flux. Implementing thermodynamic constraints prevents fluxes through these thermodynamically infeasible reactions; additionally, this limits the formation of thermodynamically infeasible loops that carry net fluxes around closed cycles in the metabolic network [217]. To implement thermodynamic constraints on the models, the minimum and maximum transformed Gibbs free energy of Recon3D reactions were obtained from the Virtual Metabolic Human database [218]. For all reactions where both the minimum and maximum ΔG values were greater than zero (and for all reversible reactions where both the minimum and maximum ΔG values were less than zero in the reverse direction), the maximum flux $v_{max,j}$ through that reaction was set to zero.

4.2.1.9 Media constraints

To simulate experimental cell culture conditions, the modeled external compartment contained all metabolites found in RPMI-1640 cell culture media (Thermo Fisher #11875) as well as fetal bovine serum (FBS) [219]. The external compartment also

contained water, hydrogen ions, hydroxide ions, oxygen, and carbon dioxide. Modeled cells were allowed to uptake these available metabolites but not any other metabolites found in Recon3D.

4.2.1.10 Code/model availability

Code for the generation and simulation of personalized FBA models is available at https://github.com/kemplab/FBA-pipeline. Personalized models can be developed for any human sample (such as cell lines or patient tumors) with RNA-seq gene expression data (and mutation data if available). Jupyter notebooks are available for 1) the processing of sample RNA-seq data to estimate enzyme abundance values; 2) the processing of sample mutation data to estimate kinetic rate parameters; 3) running FBA or FVA analysis with user-specified model constraints, media constraints, objective function, and samples of interest. TCGA models developed for this study are available as well.

4.2.2 Experimental methods

4.2.2.1 Cell culture

Table 4-3 provides the matched radiation-sensitive and radiation-resistant cell lines used for experimental validation of model predictions. All cell lines were maintained in RPMI-1640 cell culture media (Thermo Fisher #11875) with 10% fetal bovine serum (Sigma-Aldrich #F4135) at 37°C and 5% CO_2, and were free of *Mycoplasma*.

Table 4-3. Matched radiation-sensitive and radiation-resistant cell lines.

Cancer Type	Radiation-Sensitive	Radiation-Resistant	Notes	Source
Breast (BRCA)	MDA-MB-231 NQO1(-)	MDA-MB-231 NQO1(+)	Stable NQO1 expression was restored in NQO1(-) cells to create NQO1(+) cells.	Dr. David Boothman, Indiana University[220]
Glioblastoma (GBM)	M059J	M059K	Both isolated from same tumor specimen. M059J cells lack DNA-PK activity, rendering them more radiation-sensitive.	ATCC
Head and Neck (HNSC)	SCC-61	rSCC-61	rSCC-61 cells were derived from SCC-61 cells after repeated radiation exposure and selection of surviving colonies.	Dr. Cristina Furdui, Wake Forest University[175]

4.2.2.2 siRNA transfections

Cells were seeded in 24-well plates at 5×10^4 cells per well for E_{hc} GSH/GSSG measurements, or 96-well plates at 8×10^3 cells per well for glucose oxidase response measurements. 24 hours after seeding, cells were transfected using the N-TER Nanoparticle siRNA Transfection System (Sigma-Aldrich #N2913) at a final siRNA concentration of 50 nM with serum-free medium for 4 hours; afterwards, an equal volume of 2x (20%) FBS-containing medium was added to each well. Negative controls consisted of transfection with the MISSION siRNA Universal Negative Control (Sigma-Aldrich #SIC001, Lot # WDAA 1199). Predesigned siRNA's targeting individual genes were ordered from Sigma-Aldrich; the top 3 rated siRNA's based on expected efficacy for each

gene target were pooled together and transfected concurrently, except for GLUD1, GLUD2, and SUCLG2, where only 2, 1, and 1 siRNA's were available, respectively (Table 4-4). Figure 4-9 shows the knockdown efficiency of siRNA transfections in each cell line with GAPDH siRNA compared to no siRNA (just N-TER) and negative control siRNA. Experiments were performed 24hr after transfection.

Table 4-4. siRNA's used for each gene target

Gene Target	siRNA 1	siRNA 2	siRNA 3
ALDH1L1	SASI_Hs01_00106766	SASI_Hs01_00106767	SASI_Hs01_00106768
ALDH4A1	SASI_Hs01_00242114	SASI_Hs01_00242115	SASI_Hs01_00242116
ALDOA	SASI_Hs01_00211472	SASI_Hs01_00211474	SASI_Hs01_00211476
CAT	SASI_Hs02_00332471	SASI_Hs01_00092507	SASI_Hs01_00092508
CPT2	SASI_Hs01_00121386	SASI_Hs01_00121387	SASI_Hs01_00121388
DLST	SASI_Hs01_00176880	SASI_Hs01_00176881	SASI_Hs01_00176882
FH	SASI_Hs01_00037389	SASI_Hs01_00037390	SASI_Hs01_00037391
G6PD	SASI_Hs01_00013421	SASI_Hs01_00013422	SASI_Hs02_00302480
GLUD1	SASI_Hs01_00082304	SASI_Hs01_00082305	
GLUD2	SASI_Hs01_00044099		
GSR	SASI_Hs01_00152913	SASI_Hs01_00152914	SASI_Hs02_00302987
LDHB	SASI_Hs01_00181812	SASI_Hs01_00181813	SASI_Hs02_00333601
MTHFD1	SASI_Hs01_00090500	SASI_Hs01_00090501	SASI_Hs01_00090502
MTHFR	SASI_Hs01_00228224	SASI_Hs01_00228225	SASI_Hs01_00228226
NMNAT3	SASI_Hs01_00068505	SASI_Hs01_00068508	SASI_Hs01_00068509
OGDH	SASI_Hs01_00171350	SASI_Hs01_00171351	SASI_Hs01_00171353
PDHB	SASI_Hs01_00164768	SASI_Hs01_00164769	SASI_Hs01_00164770
PGAM2	SASI_Hs01_00037889	SASI_Hs02_00302252	SASI_Hs01_00037890
PGD	SASI_Hs01_00134091	SASI_Hs02_00334150	SASI_Hs01_00134092
PRDX6	SASI_Hs02_00338555	SASI_Hs01_00182363	SASI_Hs01_00182364
SHDA	SASI_Hs02_00337042	SASI_Hs01_00131081	SASI_Hs02_00337043
SUCLG2	SASI_Hs01_00106317		
TKT	SASI_Hs01_00241972	SASI_Hs01_00241973	SASI_Hs01_00241971
TPI1	SASI_Hs01_00014384	SASI_Hs01_00014385	SASI_Hs01_00014386

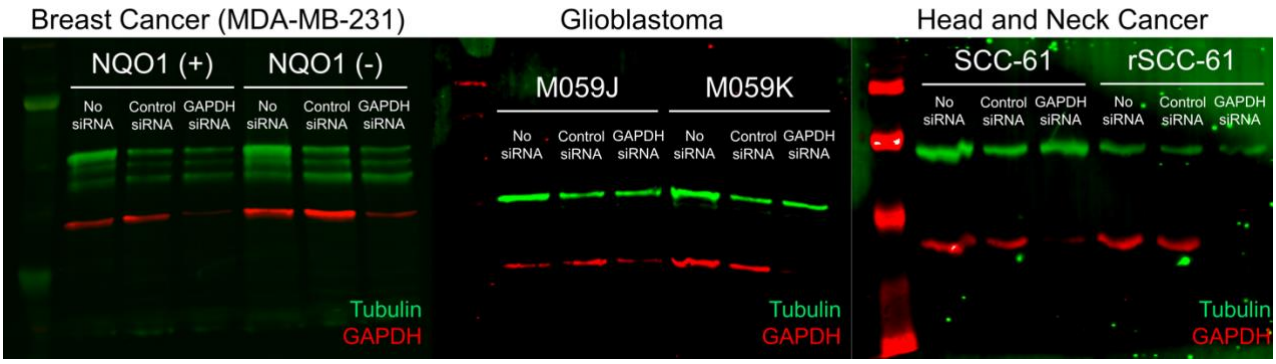


Figure 4-9. Knockdown efficiency of siRNA transfections in each cell line with GAPDH (red) siRNA compared to no siRNA (just N-TER) and negative control siRNA. Tubulin (green) was used as a loading control.

4.2.2.3 E_{hc} GSH/GSSG measurements

The protocol and reagents used for measurement of glutathione half-cell potential were adapted from [221]. KPE buffer was made by combining 16mL of solution A (6.8g KH_2PO_4 in 500mL dH_2O) and 84mL of solution B (8.5g K_2HPO_4 in 500mL dH_2O). After removal of cell culture media, 150µL of extraction buffer (0.1% Trition-X100 and 0.6% sulfosalicylic acid in KPE) to each well of a 24-well plate containing samples of interest. Plates were shaken at 800rpm for 10min to promote cell lysis. For each well, 100µL of lysate was taken for GSSG measurement, and 20µL was taken for GSSG+GSH measurement.

For GSSG+GSH measurement, 20µL samples were placed in a clear 96-well plate. 20µL of GSSG and GSH standards at concentrations of 2e-2, 1e-2, 5e-3, 2.5e-3, 1e-3, 5e-4, 2.5e-4, 1e-4, 5e-5, 2.5e-5, 1e-5, and 0 mg/mL were placed in separate rows of the 96-well plate. 120µL of 1:1 DTNB:GR solution (DTNB: 2mg 5,5'-dithiobis(2-nitrobenzoic acid) in 3mL KPE; GR: 40µL of glutathione reductase enzyme (250 units/mL) in 3mL

KPE) was added to each well. 30sec later, 60μL of NADPH solution (2mg of β-NADPH in 3mL KPE) was added to each well. Immediately after, absorbance at 412nm in each well was measured every 31sec for a period of 5min10sec (11 measurements) using a BioTek Synergy 4 plate reader. All sample and standard values were background subtracted. To determine the concentration of GSSG+GSH in each sample (mg/mL), the slope of absorbance vs. time from each sample was compared to the average slope of the GSSG and GSH standard concentrations.

For GSSG measurement, 100μL samples as well as 100μL of GSSG and GSH standards were placed in a 96-well plate. 10μL of 2VP solution (2-vinylpyridine diluted 1:50 in KPE) was added to each well. 60min later, 10μL of triethanolamine solution (triethanolamine diluted 1:10 in KPE) was added to each well. 10min later, 20μL from each well was transferred to a new clear 96-well plate. GSSG was then measured analogously to GSSG+GSH. GSSG measurements were multiplied by 1.2 to account for the dilution of the original 100μL sample with 10μL of 2VP solution and 10μL of triethanolamine solution.

To determine the concentration of GSH (mg/mL) for each sample, the measured concentration of GSSG was subtracted from the measured concentration of GSSG+GSH. To convert lysate concentrations (mg/mL) to intracellular concentrations (M), the following conversion was used:

$$[M] = \frac{[mg/mL] \times V_{sample}}{MM \times V_{cell} \times n_{cells}} \tag{25}$$

where MM is the molar mass of either GSH or GSSG (g/mol), V_{sample} is the volume of sample (20μL), V_{cell} is the estimated volume of each cell (1 pL), and n_{cells} is the number of cells per sample ($5 \times 10_4$).

The glutathione half-cell potential E_{hc} is calculated using the Nernst equation:

$$E_{hc} = E_{hc}^{o} - \frac{RT}{zF}\ln\left(\frac{[\mathrm{GSH}]^2}{[\mathrm{GSSG}]}\right) \quad (26)$$

where E_{hco} is the standard half-cell potential at pH 7.4 (-264 mV), R is the universal gas constant (8.314 J/K/mol), T is the temperature (310.15 K), z is the number of electrons (2), and F is the Faraday constant (96485 C/mol = 96.485 J/mV/mol).

4.2.2.4 Glucose Oxidase Response Measurements

Prior to seeding, cells were stained with CellTracker Red CMTPX Dye (ThermoFisher Scientific, C34552) for 30min. $8 \times 10_3$ cells per well were seeded in white-sided clear-bottom 96-well plates. 24hr after transfection and prior to treatment with glucose oxidase, CellTracker fluorescence measurements in each well were taken at Ex 577nm, Em 602nm using a BioTek Synergy 4 plate reader as a proxy for number of cells per well. An average of 10 fluorescence measurements was taken, and measurements were background subtracted. Samples were then treated with 10 mU/mL glucose oxidase enzyme (Millipore Sigma, G7141) diluted in 100µL cell culture medium (RPMI-1640 + 10% FBS) for 2hr. Afterwards, 100µL of CellTiter-Glo 2.0 Cell Viability Assay (Promega, G9241) was added to each well. Plates were shaken at 800rpm for 2min to promote cell lysis, and then incubated at room temperature for 10min. Luminescence measurements were then taken using a BioTek Synergy 4 plate reader and background subtracted. Cell viability normalized by cell count was measured by dividing the luminescence measurement by the average CellTracker fluorescence measurement.

4.3 Results

4.3.1 Automated bioinformatics pipeline for integrating multi-omic data

The framework for building FBA models of tumor metabolism was initiated with the community-curated Recon3D human metabolic reconstruction (8,401 metabolites, 13,547 reactions, and 3,268 genes) [7]. The stoichiometric representation of this reconstruction, along with minimum and maximum constraints on reaction fluxes, results in a solution space for the steady state fluxes throughout the metabolic network. Maximizing an objective function narrows the possible solution space and allows for assessment of metabolic phenotypes between tumor models. For example, using the objective function of maximizing reduction of $NADP^+$ to NADPH provides estimates of maximum cellular NADPH production. Additionally, flux variance analysis (FVA) predicts minimum and maximum fluxes through individual metabolic reactions while satisfying the maximization of the original objective function; e.g. fluxes through NADPH-generating reactions while maintaining maximum overall NADPH production. Under the assumption that tumors favor maximizing the production of reduced redox cofactors such as NADPH to decrease ROS-mediated damage induced by ionizing radiation, the associated metabolic fluxes resulting from maximizing cofactor reduction provides accurate predictions of tumor metabolism in response to radiation therapy.

To generate FBA models that are personalized to individual radiation-sensitive and radiation-resistant TCGA patient tumors, a Michaelis-Menten V_{max} constraint was set on all 4,367 Recon3D reactions with both a gene-protein-reaction (GPR) rule and enzyme commission (EC) number (Figure 4-10). This constraint sets the maximum flux for each reaction (V_{max}, units of mmol gDW^{-1} hr^{-1}) equal to the kinetic rate constant of the enzyme

catalyzing the reaction (k_{cat}, units of hr^{-1}), multiplied by the estimated protein abundance of the enzyme ([E], units of mmol gDW^{-1}). A custom protein prediction pipeline was developed to convert RNA-seq gene expression data from individual TCGA samples into estimated enzyme abundances with quantitative accuracy (Figure 4-2). This pipeline takes advantage of models developed by *Schwanhäusser et al.* relating the mRNA and protein abundances for individual genes/proteins using experimentally-measured transcription, translation, and degradation rates [212]. Additionally, a custom kinetic rate pipeline was created to extract turnover numbers for Recon3D metabolic reactions from experimentally-measured values in the BRENDA database (Figure 4-8) [137]. Envision database scores were applied to predict the effect of mutations in individual TCGA samples on the catalytic rate of corresponding metabolic enzymes [222]. Finally, standard transformed changes in Gibbs free energy ($\Delta G'^{\circ}$) from the Virtual Metabolic Human (VMH) database were used to set thermodynamic constraints, where only reactions with a negative $\Delta G'^{\circ}$ can carry non-zero net fluxes [218]. These proteomic, kinetic, and thermodynamic constraints allow for the prediction of quantitative, patient-specific, and physiologically-accurate metabolic fluxes on a genome scale.

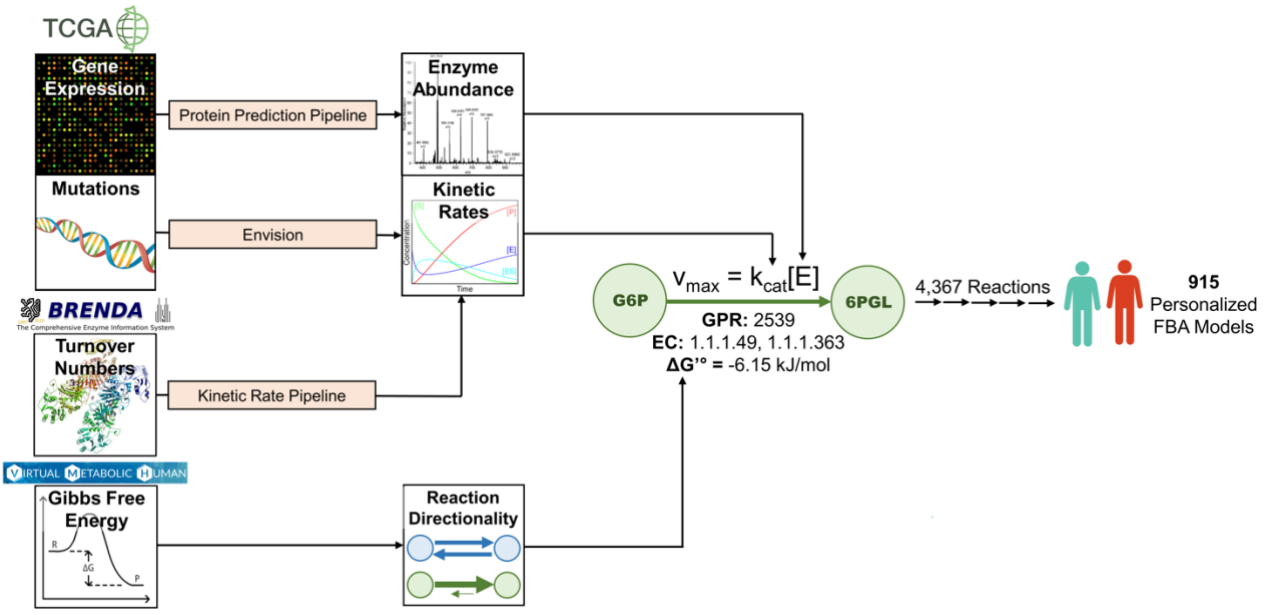


Figure 4-10. Pipeline for integration of multi-omic data into FBA models. Gene expression data from individual tumors was used to estimate the abundances of metabolic enzymes. Experimental turnover numbers from the BRENDA database and mutation data from individual tumors were combined to estimate the kinetic rates of metabolic enzymes. Together, these were used to impose a Michaelis-Menten V_{max} constraint ($V_{max} = k_{cat}[E]$) on all reactions with an available GPR rule (denoting the Entrez ID's of all genes encoding the enzyme which catalyzes the reaction) and EC number (denoting the metabolic function of the enzyme). Additionally, standard transformed Gibbs free energies from the VMH database were used to set the directionalities of metabolic reactions.

RECIST classification of TCGA samples provided an evaluation metric of radiation sensitivity based on changes in tumor size in response to radiation therapy (Figure 4-11). Patients with a complete or partial response to radiation (greater than 30% decrease in tumor size) were classified as radiation-sensitive, and patients with stable or progressive

disease (either less than 30% decrease in tumor size, or increase in tumor size) were classified as radiation-resistant. If a patient received multiple courses of radiation therapy, they were classified based on the response to their first course. Using this classification, 716 personalized FBA models of radiation-sensitive patients and 199 personalized models of radiation-resistant patients were generated.

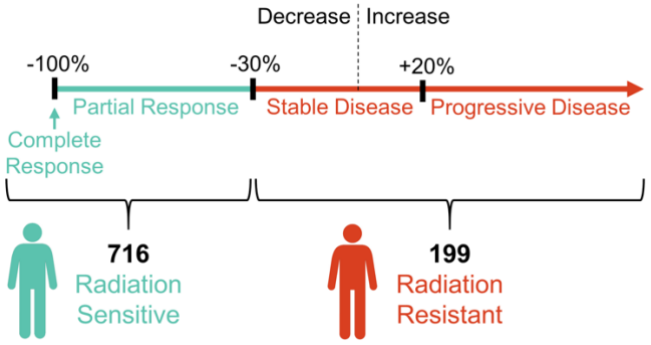


Figure 4-11. Classification of TCGA patient tumors into radiation-sensitive and radiation-resistant groups from recorded decreases/increases in tumor size after radiation therapy.

4.3.2 Compartmental differences in redox metabolic fluxes

Personalized FBA models of TCGA samples were first used to predict the capabilities of tumor cells to increase their cellular stores of reduced redox couples, which would in turn increase their antioxidative capacity and reduce ROS-mediated damage from ionizing radiation [51, 207]. Radiation-resistant tumor models were capable of producing significantly elevated levels of reduced cofactors NADPH, NADH, and glutathione (Figure 4-12). To validate model predictions, we performed experimental measurements in matched pairs of radiation-sensitive and -resistant cell lines across three different cancer types (Table 4-3). Two of the three radiation-resistant cell lines had significantly more reduced glutathione half-cell potentials than their matched radiation-sensitive cell line;

additionally, the BRCA cell line pair with a more reduced E_{hc} in the radiation-sensitive cell line had a much smaller observed difference compared to the other two pairs (Figure 4-13).

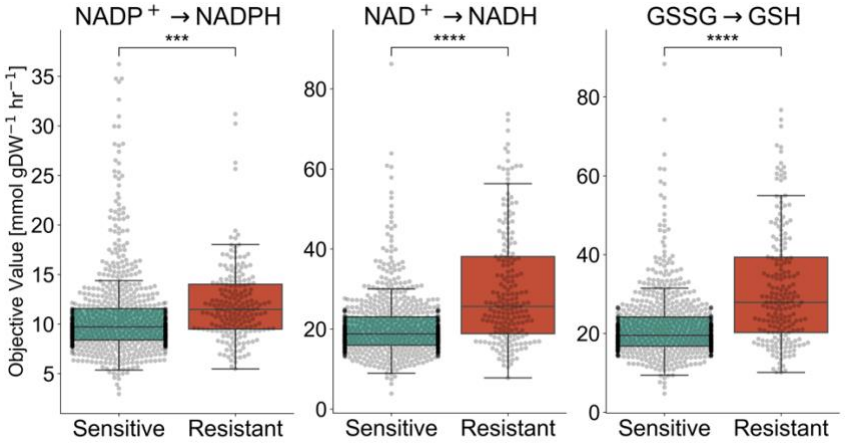


Figure 4-12. FBA-predicted production of reduced NADPH, NADH, and glutathione (GSH) in TCGA tumors.

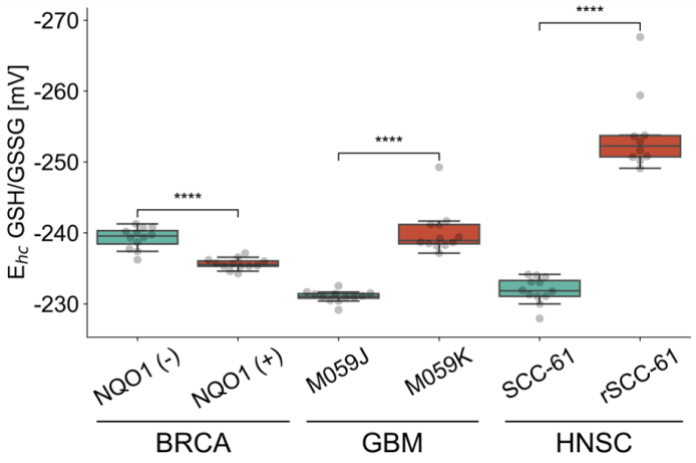


Figure 4-13. Experimentally-measured glutathione half-cell potential in cancer cell lines.

While the total cellular production of NADPH was elevated in radiation-resistant tumor models, production of NADPH in the cytosol was not increased; interestingly, mitochondrial NADPH production accounted for the differences seen between radiation-sensitive and -resistant tumors (Figure 4-14). Mitochondrial stores of NAD(P)H are required by enzymes localized to the mitochondria, many of which are involved in nucleotide metabolism, including dihydrofolate reductase and dihydroorotate dehydrogenase [223, 224]. In agreement with this observation, radiation-resistant tumor models were capable of producing increased levels of deoxynucleotides, suggesting increased capacity for DNA-damage repair following exposure to ionizing radiation (Figure 4-15) [201]. On the other hand, fatty acid production relies heavily on cytosolic stores of NAD(P)H; FBA models indicate no significant differences in production of fatty acid precursors including palmitate, suggesting a lesser role of cytosolic NADPH in differentiating metabolic phenotypes between radiation-sensitive and radiation-resistant tumors (Figure 4-15) [58].

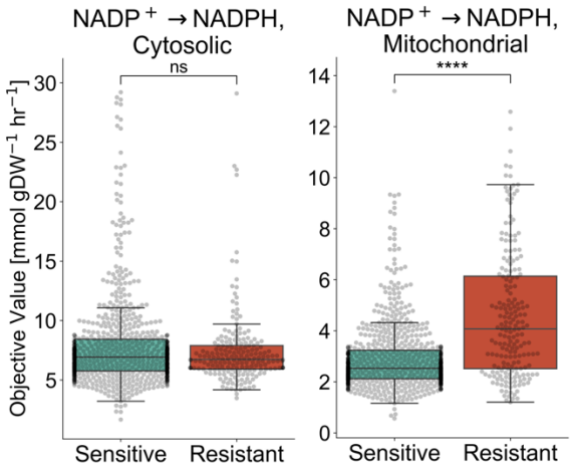


Figure 4-14. Comparison of FBA-predicted cytosolic and mitochondrial production of reduced NADPH.

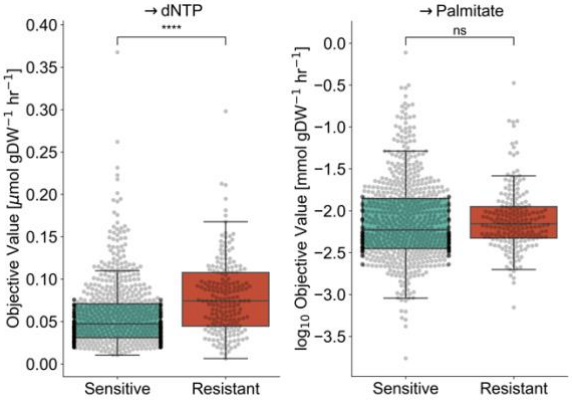


Figure 4-15. Downstream effects of increased mitochondrial but not cytosolic NADPH production in radiation-resistant tumor models. (Left) Overall production of deoxynucleotides. (Right) Overall palmitate production.

Flux variance analysis (FVA) was used to compare the average fluxes through major NADPH-generating metabolic reactions between radiation-sensitive and -resistant tumors (Figure 4-16). Surprisingly, predicted fluxes through major cytosolic reactions including

G6PD and PGD were greater in radiation-sensitive tumors compared to -resistant tumors; this observation is in agreement with previously identified flux distributions in matched radiation-sensitive and -resistant head and neck cancer cell lines [206]. On the other hand, predicted fluxes through major mitochondrial reactions including GLUD1/2, ME2/3, NNT, IDH2, and MTHFD2 were greater in radiation-resistant tumors, allowing these tumors to maximize mitochondrial NADPH stores. Hierarchical clustering of NADPH-producing fluxes yielded pronounced separation of cytosolic and mitochondrial reactions except for the folate-associated reactions MTHFD1 and MTHFD2, both of which displayed increased fluxes in radiation-resistant tumor models (Figure 4-17). The majority of radiation-resistant tumors formed a few distinct clusters characterized by increased fluxes through mitochondrial and folate-dependent NADPH-producing reactions. Clustering of samples based on radiation response was found to be significantly greater than based on other clinical factors including cancer type, as measured by the silhouette coefficient between samples (Figure 4-18). Collectively, these results demonstrate that differences in mitochondrial NADPH metabolism can discriminate between radiation-sensitive and -resistant tumors.

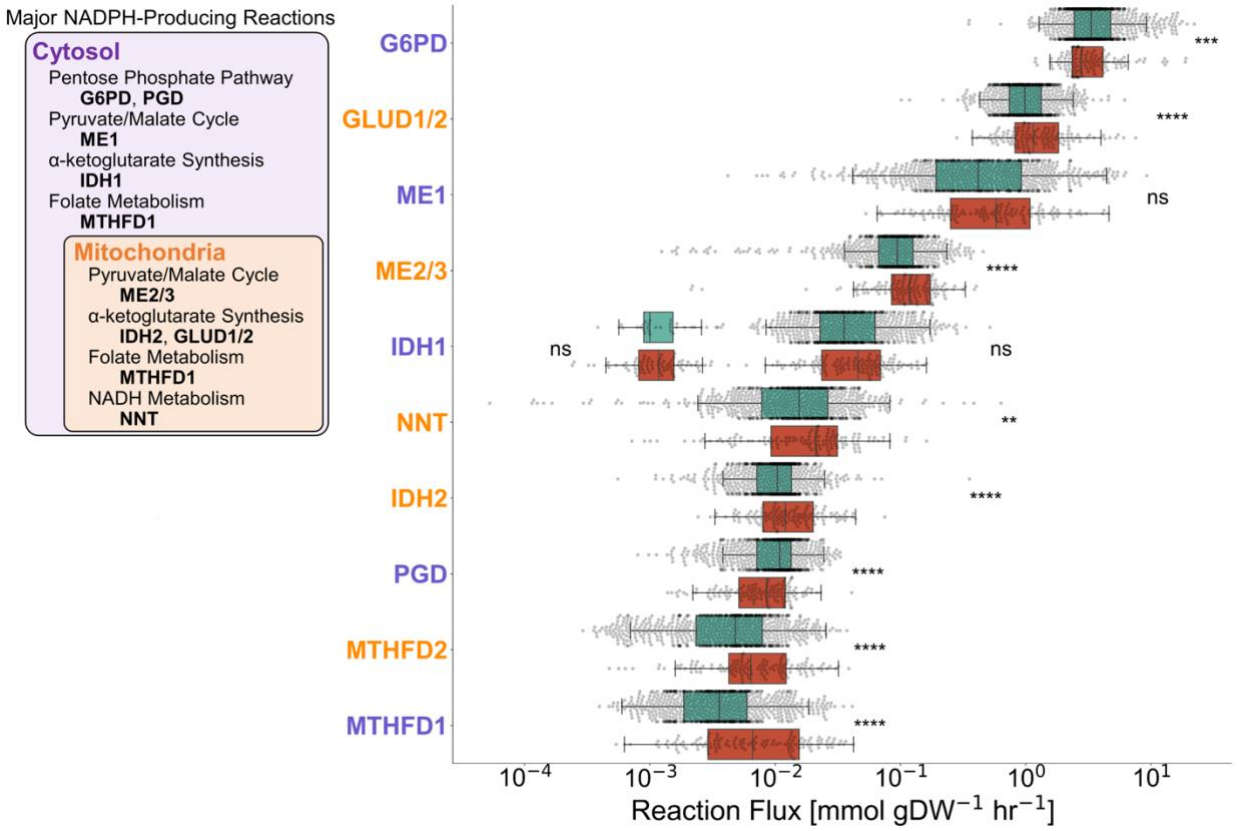


Figure 4-16. (Left) Major NADPH-producing reactions with their associated cellular compartments and metabolic pathways. (Right) FVA-predicted average fluxes through major NADPH-producing reactions. Average fluxes were calculated as the mean of minimum and maximum fluxes through each reaction while maintaining maximum total NADPH production. IDH1 fluxes were separated based on tumors with IDH1 R132 mutations (left) and wild-type IDH1 tumors (right). Reaction names are colored based on cellular compartment.

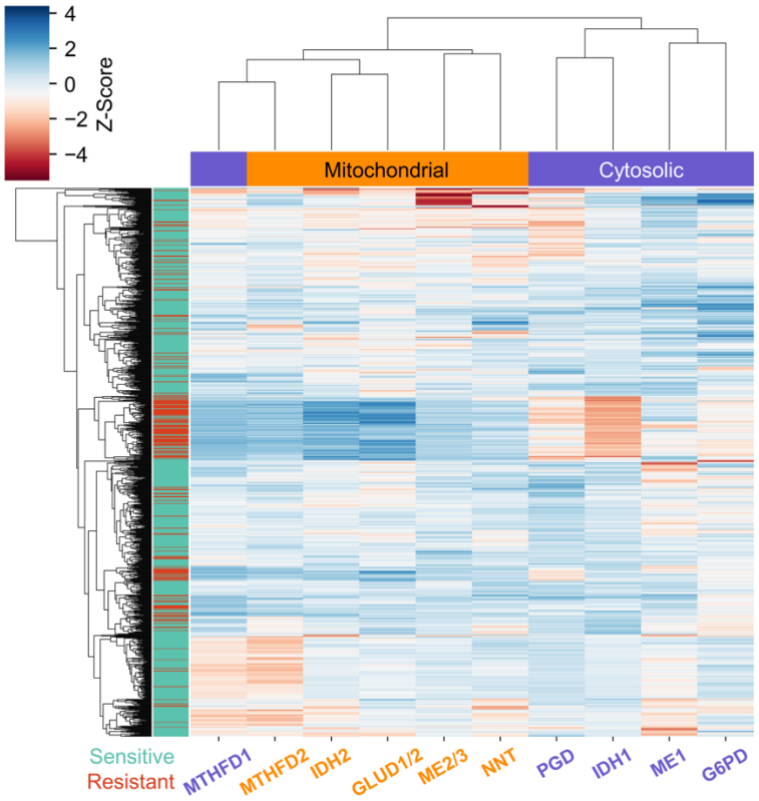


Figure 4-17. Hierarchical clustering of FVA-predicted fluxes based on TCGA patient tumor (rows) and NADPH-producing reaction (columns). Values are the Z-score of reaction fluxes across all tumors for each individual reaction.

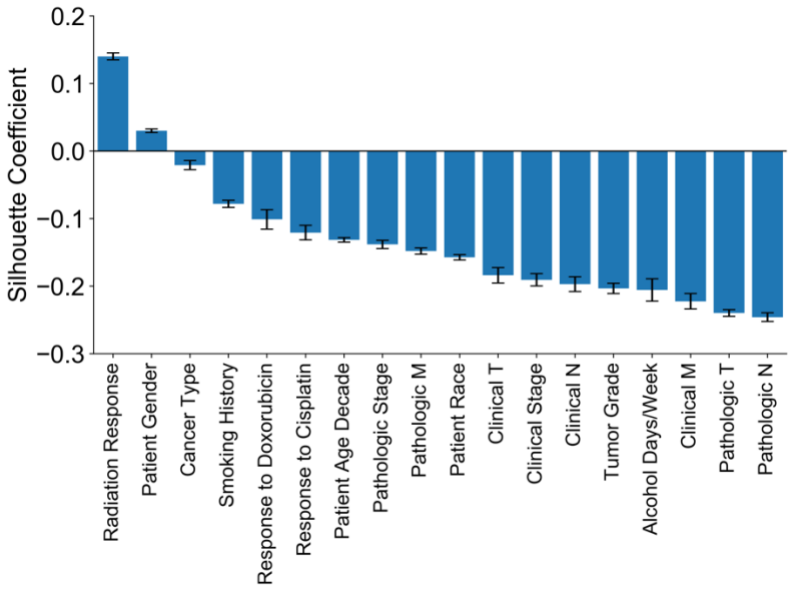


Figure 4-18. Silhouette coefficient between samples comparing radiation response and other clinical factors as cluster labels. For each factor, a more positive average silhouette coefficient signifies that samples cluster strongly together based on their factor value, and a more negative average silhouette coefficient signifies that samples cluster weakly based on their factor value (and that some other factor is likely to be the true separating factor).

4.3.3 Heterogeneity in personalized metabolic flux profiles

Although radiation-resistant tumor models displayed overall changes in redox metabolism compared to radiation-sensitive models, heterogeneity among the 199 radiation-resistant tumors may result in divergent metabolic flux profiles and thus differences in optimal therapeutic strategies for improving radiation sensitivity [225]. To determine if patient clinical factors such as cancer type and stage could be used to distinguish radiation-resistant tumor models with differing metabolic phenotypes, the correlation between these clinical factors and predicted fluxes through major NADPH-

generating reactions was determined (Figure 4-19). While cancer type, patient age, and tumor grade were highly correlated with predicted fluxes through most reactions, other factors were associated with a few select reactions. Smoking history (whether the patient was a current smoker, former smoker, or lifelong non-smoker) was highly associated with fluxes through G6PD and PGD, in agreement with previous experimental studies showing that exposure to cigarette smoke causes an upregulation of G6PD and shifts glucose metabolism towards the pentose phosphate pathway for increased NADPH production [226, 227]. Additionally, response to cisplatin treatment was highly associated with fluxes through mitochondrial reactions IDH2 and NNT; treatment with cisplatin has been previously reported to induce a mitochondrial-ROS response, which could lead to the upregulation of mitochondrial NADPH-generating enzymes [106, 228].

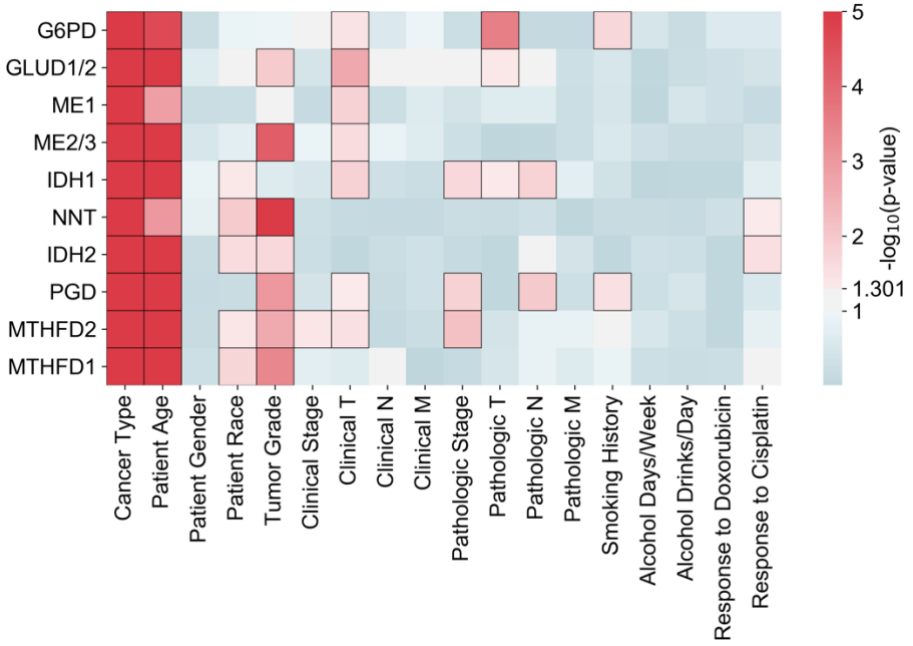
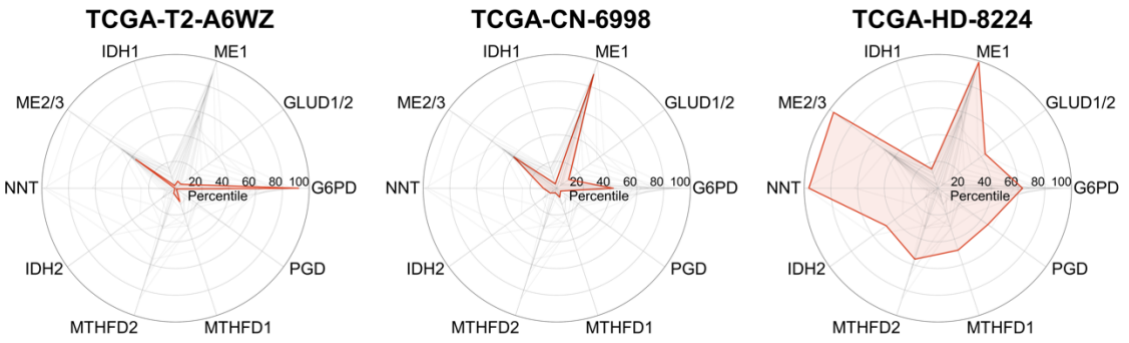


Figure 4-19. Correlation between patient clinical factors and predicted fluxes through major NADPH-producing reactions among radiation-resistant patients. Values are represented as the p-value of either the univariate regression (for

numerical factors) or 1-way ANOVA (for categorical factors) between reaction fluxes and clinical factor values. Statistically significant ($p \leq 0.05$) associations are represented with black borders.

Nonetheless, radiation-resistant patients with similar clinical factors may display stark differences in redox metabolism. Figure 4-20 illustrates personalized metabolic flux profiles of three radiation-resistant patients with head and neck squamous cell carcinoma, a cancer type characterized by high genetic and metabolic heterogeneity [229]. Although these three patients all have very similar clinical factors, the predicted fluxes through NADPH-generating reactions differ substantially; whereas some tumor models show increased NADPH generation through individual reactions (TCGA-T2-A6WZ: G6PD; TCGA-CN-6998: ME1), others show increased fluxes through multiple NADPH-generating reactions (TCGA-HD-8224: ME1, ME2/3, NNT). These findings demonstrate the utility of personalized genome-scale metabolic models for identifying targets for inhibiting NADPH production and increasing radiation sensitivity in individual patients.



	TCGA-T2-A6WZ	TCGA-CN-6998	TCGA-HD-8224
Age	53	53	63
Gender	Male	Male	Male
Race	White	White	White
Cancer Type	HNSCC	HNSCC	HNSCC
Location	Oral Tongue	Oral Tongue	Base of Tongue
Tumor Grade	G2	G2	G2
Clinical Stage	Stage IVA	Stage IVA	Stage III
HPV Status	Negative	Negative	Negative
Smoking History	Current, 40 Pack Years	Current, 40 Pack Years	Lifelong Non-Smoker
Alcohol History	7 days/week, 3.5 drinks/day	None	7 days/week, 5 drinks/day
Cetuximab Response	Progressive Disease	N.A.	Progressive Disease

Figure 4-20. Personalized NADPH-generating flux profiles of three radiation-resistant HNSCC patients, with their associated clinical factors. Radar chart values are the percentiles of reaction fluxes across all radiation-resistant tumors for each individual reaction. Patient profiles (red, filled) are shown overlaid on top of the profiles of all other radiation-resistant HNSCC profiles (black, not filled).

4.3.4 Impact of IDH1 R132 mutations on glioma NADPH production

Mutations in the NADPH-generating enzyme isocitrate dehydrogenase (IDH1) at amino acid position R132 are common in low grade gliomas and secondary high grade gliomas, classes of tumors that are commonly resistant to radiation therapy [230, 231]. These mutations affect IDH1 activity by both inhibiting the wild type enzyme's normal function of reducing $NADP^+$ to NADPH, and by causing the mutated enzyme to catalyze a neomorphic reaction involving the oxidation of NADPH to $NADP^+$ as well as the

production of the oncometabolite 2-hydroxyglutarate (Figure 4-21) [215]. Thus, these mutations are expected to decrease cellular NADPH production via both a lack of production and a neomorphic consumption by IDH1. Previous experimental studies have investigated whether the lack of NADPH production or presence of NADPH consumption results in a greater depletion of NADPH levels; although most evidence points to the neomorphic reaction's consumption of NADPH as the major cause of NADPH depletion, this question has yet to be resolved [81, 232].

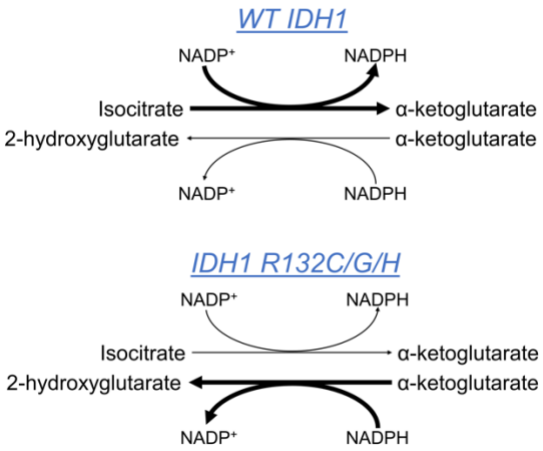


Figure 4-21. Schematic demonstrating the relative magnitudes of normal (NADPH-producing, isocitrate to α-ketoglutarate) and neomorphic (NADPH-consuming, α-ketoglutarate to 2-hydroxyglutarate) IDH1 reactions in IDH1 wild-type (top) and IDH1 R132 mutation (bottom) tumors.

Using experimentally-measured kinetic rates of wild-type and R132-mutant IDH1 enzymes, separate FBA models of 104 IDH1 R132 patients were generated by implementing (1) wild-type normal (α-ketoglutarate-producing) and wild-type neomorphic (2-hydroxyglutarate-producing) reactions; (2) mutant normal and wild-type neomorphic reactions; and (3) mutant normal and mutant neomorphic reactions (Figure 4-22) [216].

Changes in total cellular NADPH production were evaluated between models 1-2 to analyze the effect of the lack of NADPH production, and between models 2-3 to analyze the effect of the presence of NADPH consumption. Implementation of the mutant neomorphic reaction causes a substantially greater decrease in total cellular NADPH production compared to implementation of the mutant normal reaction in almost all patient tumors; thus, FBA models predict that the presence of NADPH consumption by mutant IDH1 is the major cause of NADPH depletion, in agreement with previous experimental studies. Because of the greater kinetic rate of the neomorphic NADPH-consuming reaction, NADPH production is predicted to cause the greatest impact in patients with IDH1 R132H mutations, the most common and well-characterized IDH1 mutation among gliomas, compared to R132C and R132G mutations (Figure 4-23) [216, 233]. However, the predicted decrease in cellular NADPH production due to mutant normal and neomorphic reactions is negligible compared to wild-type tumors, suggesting that IDH1 R132 mutations do not significantly alter tumor NADPH production. Despite *in vitro* experiments demonstrating that overexpression of R132-mutant IDH1 results in increased sensitivity to radiation therapy, many patient tumors with the IDH1 R132 mutation are still radiation-resistant (74.0% of patients in TCGA), possibly due to a negligible effect on NADPH production by these mutations.

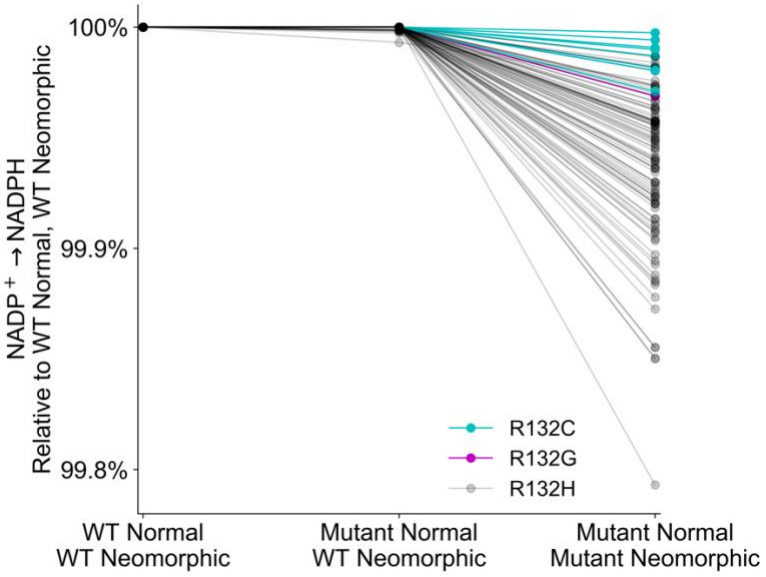


Figure 4-22. NADPH production in individual IDH1 R132C/G/H tumor models while implementing (1) wild-type normal and wild-type neomorphic reactions (left), (2) mutant normal and wild-type neomorphic reactions (middle), and (3) mutant normal and mutant neomorphic reactions (right). Values are normalized relative to NADPH production while implementing wild-type normal and wild-type neomorphic reactions.

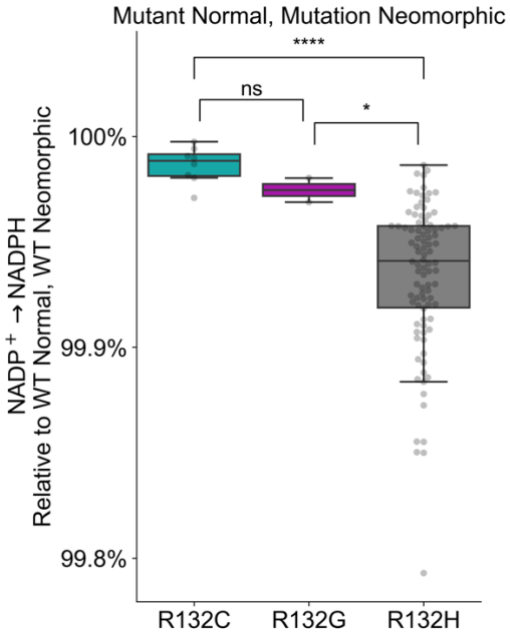


Figure 4-23. Comparison of the effect of implementing mutant normal and mutant neomorphic reactions (right column of Figure 4-22) between IDH1 R132C, R132G, and R132H tumors.

4.3.5 Simulated genome-wide knockout screen for identifying redox targets

FBA models enable the assessment of changes in metabolic phenotypes in response to knockout of a particular metabolic enzyme-encoding gene, providing insight into the gene's role in genome-scale metabolism. A simulated genome-wide knockout screen was performed to predict the effect of knocking out each individual gene in Recon3D on total cellular NADPH production across all radiation-sensitive and radiation-resistant tumor models (Figure 4-24). Most gene knockouts did not significantly decrease total NADPH production, providing evidence for the ability of metabolic networks to redirect flux through alternate, compensatory pathways following perturbation in order to optimize NADPH production [206]. Among the top knockouts across all patient tumors were genes directly involved in NADPH generation, including G6PD, GLUD1, ME1, and H6PD.

Additionally, knockout of genes involved in glycolysis (ALDOA, GAPDH, PKM, PGAM1, ENO1, HK1, GPI) resulted in significant NADPH depletion, demonstrating the strong interconnection between central carbon metabolism and NADPH generation [234]. There was also significant depletion of NADPH for genes involved in folate metabolism (DHFR), amino acid metabolism (PYCR2), and portions of pentose phosphate pathway not directly involved in NADPH production (PGLS).

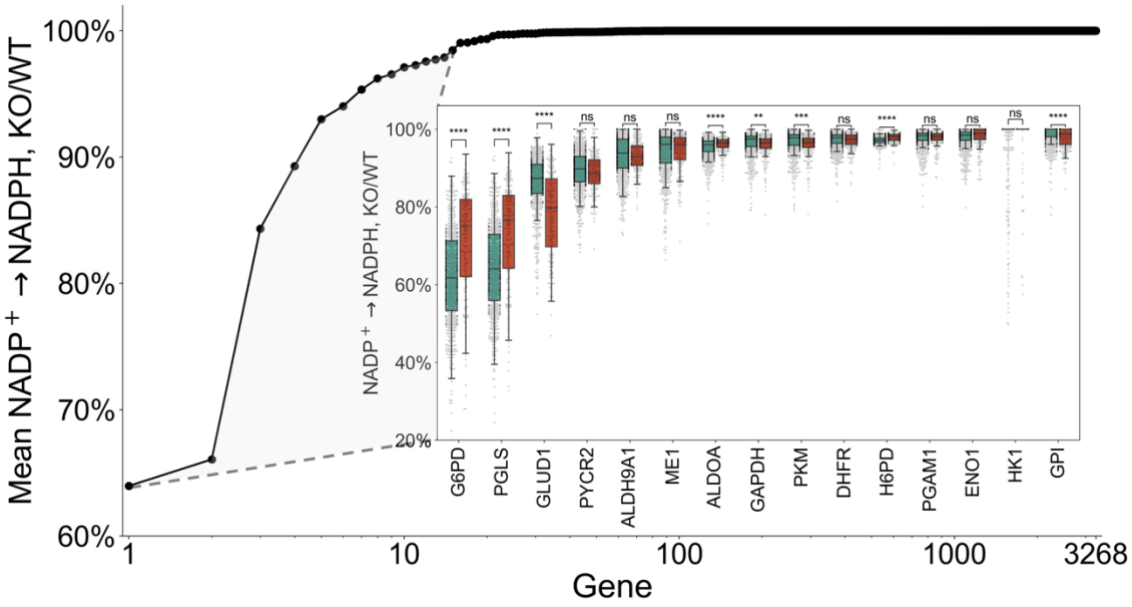


Figure 4-24. Effect of simulated knockout of each individual gene in Recon3D on total NADPH production in TCGA tumors. Values are the ratio of total NADPH production after versus before knockout. Genes are rank ordered based on increasing mean KO/WT ratio value (decreasing gene knockout effect) across all tumor models. Outset: KO/WT ratio values are averaged across all tumor models. Inset: For the top 15 genes, KO/WT ratio values from individual patient tumor

models are shown, along with the statistical significance of the comparison between radiation-sensitive and -resistant values.

A statistical comparison on the effect of each gene knockout on NADPH production was performed to determine which knockouts have the greatest differential effect in radiation-resistant versus radiation-sensitive tumors. (Figure 4-25). 26 gene knockouts caused a significantly greater decrease in NADPH production among radiation-resistant tumors compared to radiation-sensitive tumors ($\log_2$[Average Effect in Resistant Tumors / Average Effect in Sensitive Tumors] > 0; FDR-adjusted p-value < 0.05), and 24 gene knockouts caused a significantly greater decrease among radiation-sensitive tumors ($\log_2$[Resistant/Sensitive] < 0; FDR-adjusted p-value < 0.05). In agreement with observed differences in NADPH-generating reaction fluxes, knockout of genes directly involved in mitochondrial NADPH production (GLUD1/2, IDH2, ME2) and folate metabolism (MTHFD1, MTHFR) caused a greater predicted decrease in NADPH production in radiation-resistant tumors, while those involved in cytosolic NADPH production (G6PD, PGD) and other members of the pentose phosphate pathway (PGLS, RPE, RPIA, TALDO1, TKT) caused a greater decrease in radiation-sensitive tumors. Additionally, many aldehyde dehydrogenase genes (ALDH1L1, ALDH1L2, ALDH5A1) showed greater predicted decrease in NADPH production in radiation-resistant tumors, consistent with previous experimental evidence demonstrating greater susceptibility to NADPH-depleting chemotherapies [206].

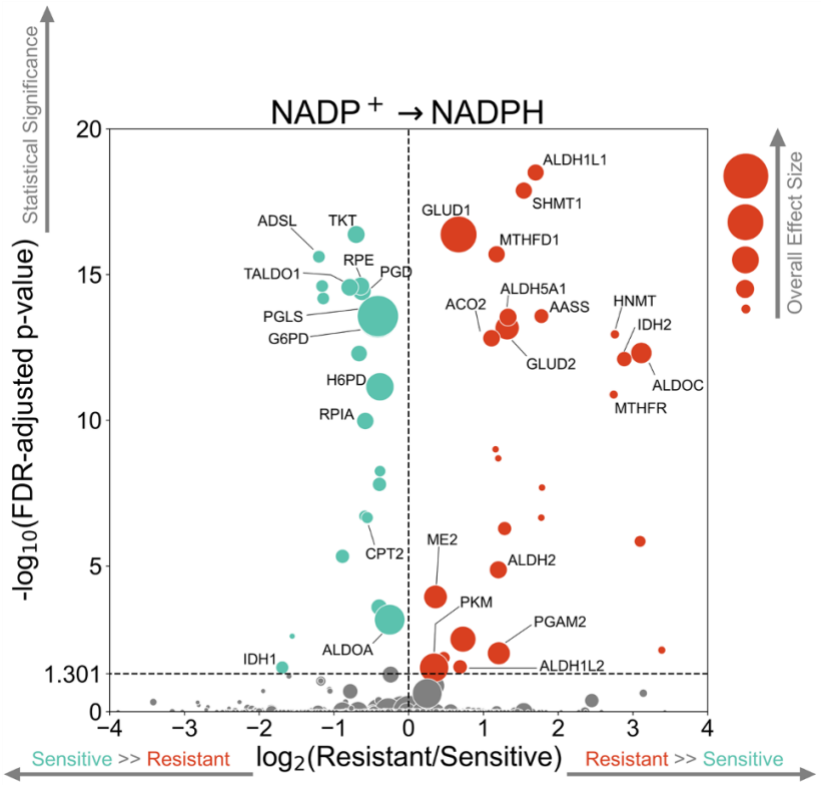


Figure 4-25. Volcano plot comparing the effect of each simulated gene knockout (individual dot) on total NADPH production between radiation-sensitive and radiation-resistant tumors. X-axis: $\log_2$(Resistant/Sensitive), where "Resistant" equals the mean (WT-KO)/WT ratio in radiation-resistant tumors, and "Sensitive" equals the mean (WT-KO)/WT ratio in radiation-sensitive tumors; values < 0 (green dot on the left of the dotted line) signifies that a knockout has a greater effect on total NADPH production in radiation-sensitive tumors, whereas values > 0 (red dot on the right of the dotted line) signifies that a knockout has a greater effect on total NADPH production in radiation-resistant tumors. Y-axis: statistical significance (false discovery rate-adjusted p-values based on the Benjamini-Hochberg procedure) of comparison of knockout effects between radiation-sensitive and radiation-resistant tumors; values above the dotted line (FDR-adjusted p-value ≤

0.05) are statistically-significant. The size of each dot is proportional to the overall effect size (mean (WT-KO)/WT ratio across all tumor models regardless of radiation sensitivity).

A similar analysis was performed to predict gene knockouts with significant effects on glutathione reduction (Figure 4-26**,** Figure 4-27). Although many genes showed overlapping effects between their impact on NADPH reduction and glutathione reduction, there were some notable differences. For instance, folate metabolism genes MTHFD1 and MTHFR were predicted to have greater impacts on glutathione reduction in radiation-sensitive tumors. Additionally, many pentose phosphate genes including G6PD, PGD, and members of the non-oxidative pathway were not predicted to significantly affect glutathione reduction. New gene targets arose having significant impacts on glutathione reduction, including genes involved in the TCA cycle (DLST, FH, OGDH, SDHA, SUCLG2), oxidation-reduction reactions (AKR1A1, GSR, PRDX6), and glycine metabolism (AMT, GCSH, GLDC), which have been previously shown to impact GSH/GSSG ratios and ROS levels in cancer [235].

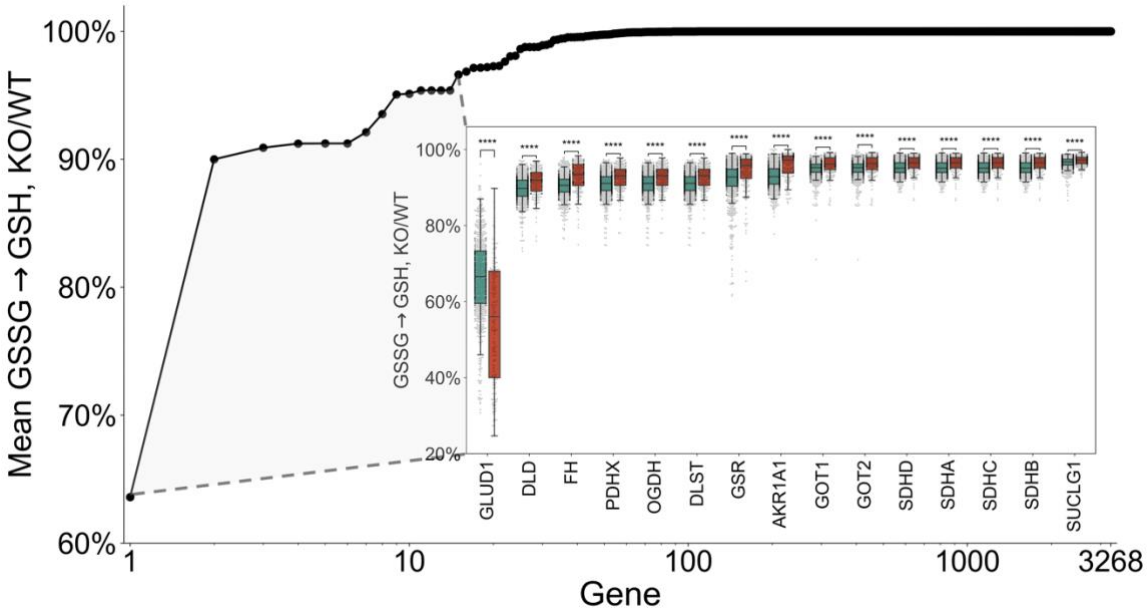


Figure 4-26. Effect of simulated knockout of each individual gene in Recon3D on total GSH production. Values are the ratio of total GSH production after versus before knockout. Genes are rank ordered based on increasing mean KO/WT ratio value (decreasing gene knockout effect) across all tumor models. Outset: KO/WT ratio values are averaged across all tumor models. Inset: For the top 15 genes, KO/WT ratio values from individual patient tumor models are shown, along with the statistical significance of the comparison between radiation-sensitive and -resistant values.

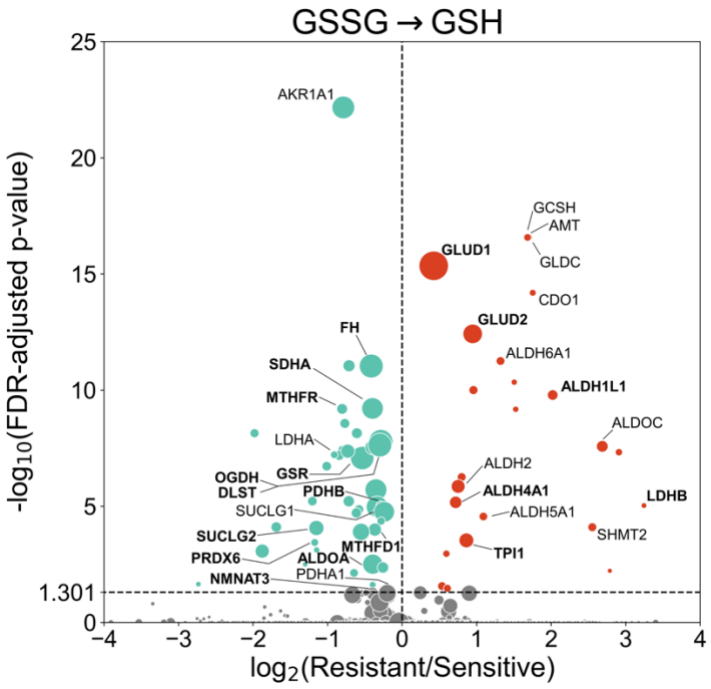


Figure 4-27. Volcano plot comparing the effect of each simulated gene knockout on total reduced glutathione production between radiation-sensitive and radiation-resistant tumors. Genes tested by experimental siRNA knockdown studies are bolded.

To validate model-predicted targets of glutathione metabolism, the change in glutathione half-cell potential (ΔE_{hc} GSH/GSSG) between siRNA knockdowns and negative control siRNA-transfected cells was measured in both radiation-sensitive and -resistant cell lines for 18 different gene targets (Figure 4-28). The difference in ΔE_{hc} between -resistant and -sensitive cell lines (ΔE_{hc}Resistant - ΔE_{hc}Sensitive) is greater than 0 if the gene knockdown causes greater oxidation in the radiation-resistant cell line, corresponding to a model-predicted $\log_2$(Resistant/Sensitive) greater than 0. The t-statistic of the 1-sample t-test comparing the three experimentally-measured values of ΔE_{hc}Res - ΔE_{hc}Sens to the null hypothesis population mean of zero (equal effect in radiation-sensitive and -resistant cell lines) also has the same sign as model-predicted $\log_2$(Resistant/Sensitive)

if model and experiment are in agreement. 4/6 model-predicted radiation-resistant targets also had more oxidative effects in the radiation-resistant cell lines; the other two targets (TPI1 and ALDH1L1) each had one cell line pair with more oxidative effects in the radiation-resistant cell line. While some model-predicted radiation-sensitive targets had more oxidative effects in the radiation-resistant cell lines, those targets with the largest predicted differential effect (PRDX6, SUCLG2, MTHFR, and GSR) all had more oxidative effects in the radiation-sensitive cell lines. These results suggest that FBA models of TCGA tumors can accurately predict gene targets which differentially affect glutathione production in radiation-sensitive and -resistant cancers.

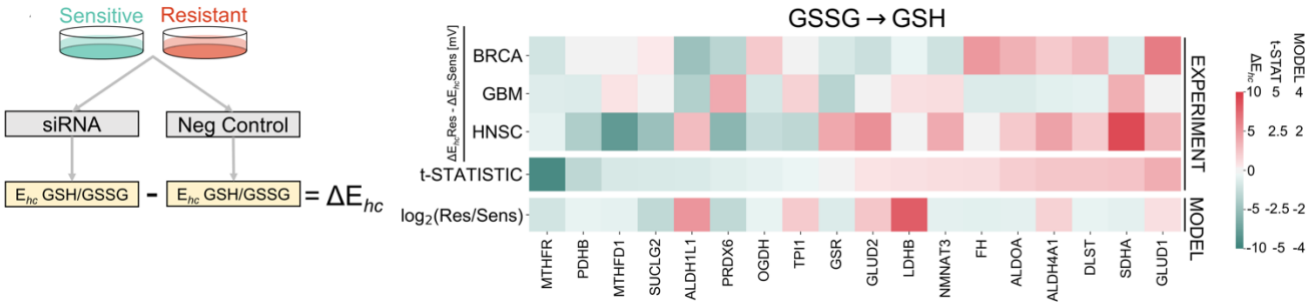


Figure 4-28. (Left) Schematic showing the measurement of ΔE_{hc} in radiation-sensitive and -resistant cancer cell lines. (Right) Comparison of model-predicted and experimentally-measured effects of gene knockdown on reduced glutathione production. Top 3 rows: ΔE_{hc}Res - ΔE_{hc}Sens in siRNA knockdowns across all three cell line pairs. Middle row: t-statistic from 1-sample t-test comparing the three experimentally-measured values of ΔE_{hc}Res - ΔE_{hc}Sens to the null hypothesis population mean of zero (equal effect in radiation-sensitive and -resistant cell lines). Bottom row: model-predicted log fold change in gene knockout effect on reduced

glutathione production between radiation-resistant and -sensitive TCGA tumor models.

4.3.6 Disparities in redox metabolism and H_2O_2-scavenging systems

Reactive oxygen species including H_2O_2 contribute significantly to the DNA damage and cellular death caused by ionizing radiation therapy [197, 198]. As expected, FBA models of radiation-resistant tumors were capable of significantly increased H_2O_2 clearance compared to radiation-sensitive tumors (Figure 4-29). To validate model predictions of response to ROS, cancer cell line pairs were treated with 10 mU/mL of glucose oxidase for 2 hours; when added to the extracellular medium with non-limiting levels of glucose, this enzyme produces hydrogen peroxide at levels that remain consistent for up to 8 hours, as compared to providing a bolus of extracellular hydrogen peroxide which can decrease to half its original concentration in as little as 15 minutes [236, 237]. Two of the three radiation-resistant cell lines showed a lesser decrease in cell viability when treated with 10 mU/mL glucose oxidase compared to their matched -sensitive cell lines, while the other pair showed no significant difference (Figure 4-30). The differential effect of glucose oxidase within the cell line pairs matches very well with the observed differences in glutathione half-cell potential, suggesting that reduced glutathione stores are utilized for H_2O_2 clearance in these cell lines.

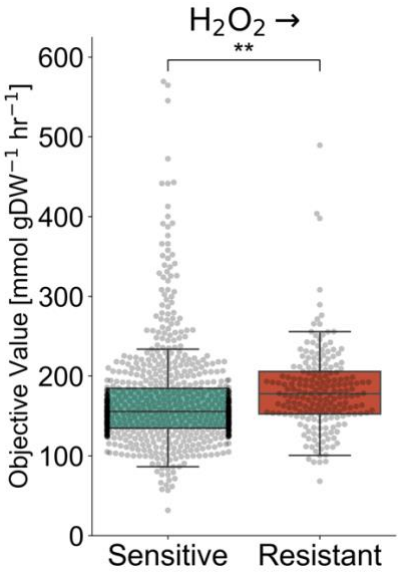


Figure 4-29. Model-predicted total clearance of H_2O_2 in TCGA tumors.

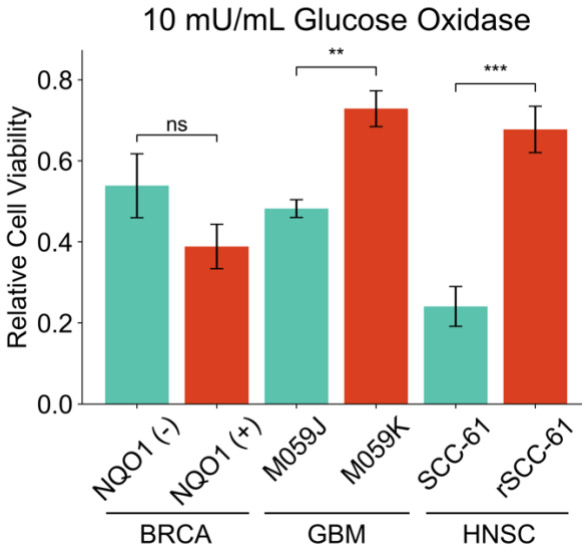


Figure 4-30. Experimentally-measured response of radiation-sensitive and -resistant cancer cell lines to 10 mU/mL of glucose oxidase, calculated as the relative cell viability compared to 0 mU/mL glucose oxidase.

To determine the most important contributors to the increased H_2O_2 clearance observed in radiation-resistant tumor models, an assessment of the major cellular H_2O_2-

clearing systems, including catalase (CAT), glutathione peroxidase (GPx), glutaredoxins (Grx), peroxiredoxins (Prx), and thioredoxins (Trx), was performed (Figure 4-31). Many of these systems utilize varying isoforms in different cellular compartments, leading to compartmental differences in enzyme abundances. Additionally, while the same CAT isoform is found in many different cellular compartments, compartmental pH differences lead to differences in observed catalytic rates. Finally, cellular stores of NADPH and glutathione cofactors that these systems rely on vary significantly between compartments, imposing limits on the maximum fluxes through these metabolic enzymes. All of these important considerations are taken into account with our bioinformatics and modeling pipeline, as opposed to more traditional modeling approaches such as small-scale kinetic models and genome-scale models without kinetic constraints. The majority of H_2O_2-clearing fluxes predicted from flux variance analysis did not differ significantly between radiation-sensitive and -resistant tumors. However, 3 of the mitochondrial H_2O_2-clearing fluxes (CAT, GPx, Grx) showed significantly greater fluxes in radiation-resistant tumors, accounting for the increased overall capacity for H_2O_2 clearance. This is in agreement with previous findings of increased capacity for mitochondrial NADPH reduction as opposed to cytosolic NADPH reduction in radiation-resistant tumor models (Figure 4-14).

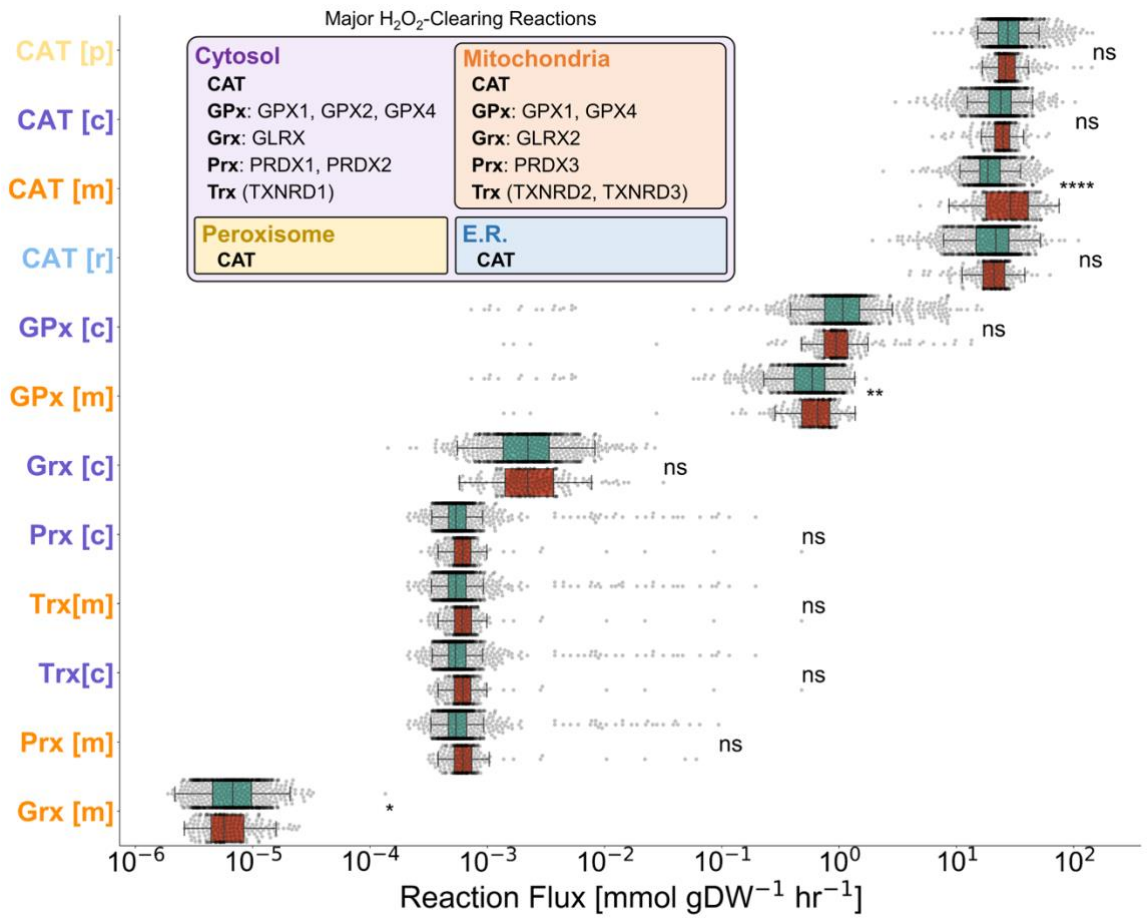


Figure 4-31. FVA-predicted average fluxes through major H_2O_2-clearing reactions. Average fluxes were calculated as the mean of minimum and maximum fluxes through each reaction while maintaining maximum total H_2O_2 clearance. Inset: Major H_2O_2-clearing reactions with their associated cellular compartments and compartment-specific isoforms.

A simulated genome-wide knockout screen was then performed to identify additional targets affecting overall H_2O_2 clearance in radiation-resistant tumors (Figure 4-32). Interestingly, while few genes directly involved in H_2O_2 clearance were found to be significant targets (except CAT), many of the same targets from the NADPH knockout screen were significant targets with similar differential effects between radiation-sensitive

and -resistant tumors. Knockout of cytosolic NADPH production genes (G6PD, PGD) and other members of the pentose phosphate pathway (PGLS, RPE, RPIA, TALDO1, TKT) showed significantly greater effects in radiation-sensitive tumor models, whereas knockout of genes involved in mitochondrial NADPH production (GLUD1/2), central carbon metabolism (ACO2, ALDOC, GAPDH, PGAM2, PKM), and folate metabolism (MTHFD1, MTHFD2, MTHFR) showed greater effects in radiation-resistant tumor models. To validate these targets, Bliss independence scores showing the effect of siRNA gene knockdown on glucose oxidase response were measured in radiation-sensitive and -resistant cell lines for 9 different gene targets (Figure 4-33). The difference in Bliss score between -sensitive and -resistant cell lines ($Bliss_{Sensitive}$ - $Bliss_{Resistant}$) is greater than 0 if the gene knockdown causes a greater decrease in cell viability with glucose oxidase treatment in the radiation-resistant cell line, corresponding to a model-predicted $\log_2$(Resistant/Sensitive) greater than 0. The t-statistic of the 1-sample t-test comparing the three experimentally-measured values of $Bliss_{Sensitive}$ - $Bliss_{Resistant}$ to the null hypothesis population mean of zero (equal effect in radiation-sensitive and -resistant cell lines) also has the same sign as model-predicted $\log_2$(Resistant/Sensitive) if model and experiment are in agreement. Although all but one (ALDH1L1) gene knockdown resulted in a greater response to glucose oxidase in the radiation-resistant cell lines, three of the model-predicted radiation-resistant targets (GLUD1, MTHFD1, and PGAM2) had the largest differential response across all siRNA's tested. Additionally, each of the four radiation-sensitive targets (CPT2, G6PD, PGD< and TKT) had a greater response in at least one of the radiation-sensitive cell lines. Interestingly, TCGA models accurately predicted that MTHFD1 knockdown would have a greater effect on glutathione reduction in radiation-

sensitive cancers but a greater effect on H_2O_2 response in radiation-resistant cancers, suggesting that FBA models accurately capture other metabolic systems which impact ROS clearance besides glutathione-dependent pathways. Overall, the agreement between model predictions and experimental validation demonstrates the ability of FBA models to correctly identify novel targets of redox metabolism which differentially impact radiation-resistant cancers.

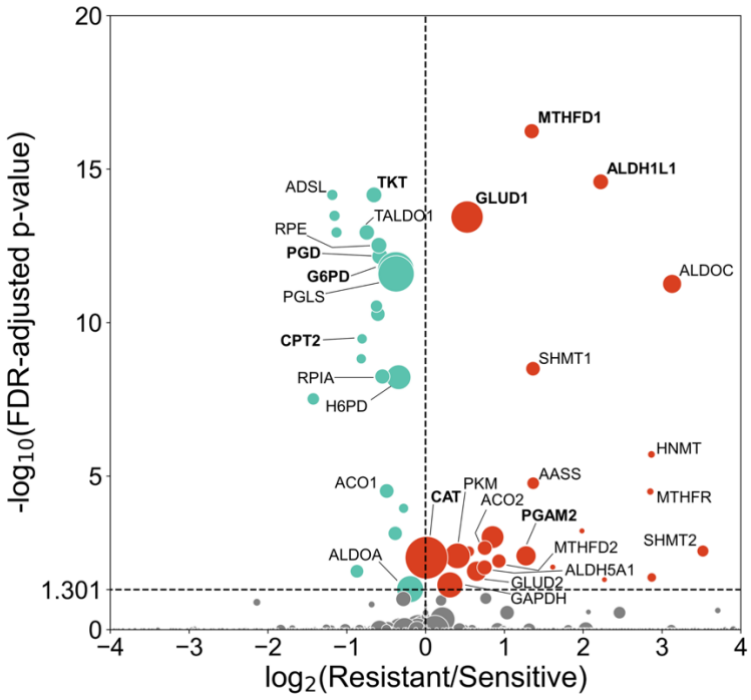


Figure 4-32. Volcano plot comparing the effect of each gene knockout on total H_2O_2 clearance between radiation-sensitive and radiation-resistant tumors. Genes tested by experimental siRNA knockdown studies are bolded.

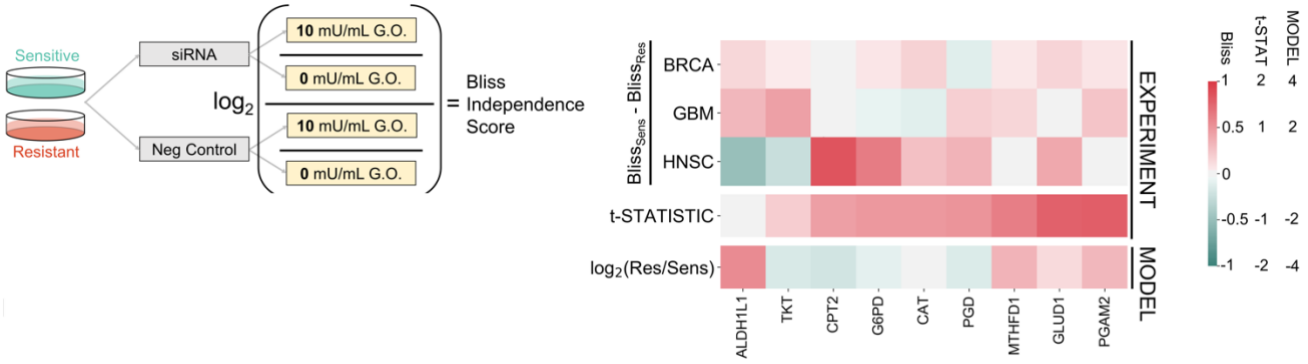


Figure 4-33. (Left) Schematic showing the measurement of Bliss independence scores in radiation-sensitive and -resistant cancer cell lines. (Right) Comparison of model-predicted and experimentally-measured effects of gene knockdown on H_2O_2 clearance. Top 3 rows: $Bliss_{Sens}$ - $Bliss_{Res}$ in siRNA knockdowns across all three cell line pairs. Middle row: t-statistic from 1-sample t-test comparing the three experimentally-measured values of $Bliss_{Sens}$ - $Bliss_{Res}$ to the null hypothesis population mean of zero (equal effect in radiation-sensitive and -resistant cell lines). Bottom row: model-predicted log fold change in gene knockout effect on H_2O_2 clearance between radiation-resistant and -sensitive TCGA tumor models.

4.4 Discussion

The role of reactive oxygen species (ROS) formation on oxidative DNA damage and ultimately cellular death during ionizing radiation treatment has been well characterized; however, the identification of targets within the redox metabolic network which may impact antioxidation, ROS response, and sensitivity to radiation therapy is still ongoing [238]. Initial studies identified G6PD as an important contributor of cellular stores of NADPH, which is used for both biosynthetic/repair processes as well as antioxidation via reduction of glutathione and thioredoxin [239, 240]. Enzymes involved in central carbon metabolism, including GLUT1, PFK, and PK were later found to significantly impact

radiation response and viewed as potential targets for improving radiation sensitivity [16, 241-244]. Systems biology studies using genome-scale datasets have recently provided unbiased approaches towards discovery of metabolic targets implicated in radiation response, including FAS, LDH, and PRDX1/2 [175]. However, our bioinformatics and metabolic modeling pipeline represents, to our knowledge, the first study to integrate multiple -omic datasets into genome-scale mechanistic models to compare redox metabolism between radiation-sensitive and radiation-resistant tumors, as well as to identify novel biomarkers which could be targeted to improve radiation therapy response.

Because redox cofactors such as NADPH are involved in nearly one thousand reactions throughout the human metabolic network, genome-scale FBA models are well-suited to study redox metabolism and their interconnections with other metabolic processes [7]. Nonetheless, few studies have used FBA models for studying cancer redox metabolism. Fan *et al*. developed an FBA model of immortalized baby mouse kidney epithelial cells to identify major contributors to the cellular NADPH pool, finding that folate-dependent enzymes MTHFD1 and MTHFD2 contributed to NADPH production as much as the more commonly-studied enzyme G6PD [144]. Although experimentally measured uptake/excretion fluxes and growth rates were used to constrain their model, most reactions were not constrained by cell type-specific information such as transcriptomic data, which is commonly used to set bounds on maximum reaction fluxes [11]. Additionally, their use of parsimonious flux balance analysis (minimizing total sum of fluxes while maintaining experimentally-measured fluxes and growth rates) yields single values for each reaction flux which do not capture the biological variability in metabolic fluxes as well as flux variance analysis [245]. *Benfeitas et al*. later developed

FBA models of hepatocellular carcinomas (HCC) to characterize heterogeneity in redox metabolism between HCC tumors at differing stages of progression and within different HCC tumor clusters [246]. Using these models, the authors discovered distinct differences in NADPH production and H_2O_2 clearance between HCC tumors with high G6PD expression and those with high ALDH2 expression. However, their use of the MADE algorithm for converting transcriptomic data into upper flux bounds results in discrete reaction constraints which lack quantitative accuracy compared to using continuous enzyme abundance values. Furthermore, kinetic and thermodynamic effects on reaction constraints, which have significant impacts on metabolic fluxes, were not present in their models [12]. Our approach overcomes these methodological shortcomings by incorporating transcriptomic and mutational data from individual patient tumors, as well as genome-scale kinetic and thermodynamic parameter values, into quantitative constraints on metabolic fluxes, allowing for more accurate predictions of metabolic differences between radiation-sensitive and -resistant tumors (Figure 4-10).

Our previous FBA modeling and experimental analysis of radiation-sensitive and -resistant head and neck cancer (HNSCC) cell lines suggested that radiation-resistant cancers were capable or rerouting flux through alternative metabolic pathways in order to achieve maximal production of redox cofactors including NADPH and NADH [206, 207]. This flux rerouting was also seen in our FBA models of radiation-resistant TCGA tumors, as increased flux through mitochondrial NADPH-producing reactions allowed for increased mitochondrial stores of reduced redox cofactors as well as increased fluxes through mitochondrial H_2O_2-clearing reactions (Figure 4-14**,** Figure 4-31). While mitochondrial-specific compartmentalization of signaling and energy metabolism has been

previously identified to impact chemotherapy and radiation response, mitochondrial compartmentalization of antioxidant stores for protection against ROS-mediated damage from ionizing radiation has not been well characterized; nonetheless, mitochondrial NADPH-producing enzymes identified from this study, including GLUD1, IDH2, ME2, and NNT, have been previously implicated in regulation of oxidative stress and tumor proliferation, and their targeting for increasing radiation sensitivity should be explored [247-252]. Additionally, simulated gene knockdown screens and validation with experimental siRNA knockdown experiments suggested that genes involved in mitochondrial (GLUD1/2) and folate-dependent (MTHFD1) NADPH production, as well as central carbon metabolism genes (LDHB, PGAM) may be novel targets for inhibiting reduced glutathione production and/or H_2O_2 clearance in radiation-resistant tumors (Figure 4-28**,** Figure 4-33). On the other hand, G6PD, the most well-studied NADPH-producing reaction, actually exemplified larger fluxes in radiation-resistant tumor models and showed lower experimental effects on targeted H_2O_2 clearance in radiation-resistant cell lines compared to the aforementioned gene targets. Interestingly, the G6PD cluster identified by *Benfeitas et al.* in hepatocellular carcinomas (HCC) is highly enriched in genes identified from this current study as radiation-sensitive targets (ALDOA, G6PD, PGD, RPE, RPIA), whereas the ALDH2 cluster is highly enriched in genes identified as radiation-resistant targets (ALDH2, ALDH5A1, ALDH6A1, CAT, MTHFD1, SHMT1) [246]. Despite the fact that the authors were not analyzing redox metabolism in the context of radiation sensitivity, this correspondence between gene sets from separate analyses may suggest that there exist two fundamental tumor subtypes with distinct redox metabolic phenotypes and corresponding radiation sensitivities.

With FBA models of patient tumors, personalized metabolic flux profiles can be generated to identify optimal gene targets for clinical subsets or individual radiation-resistant patients (Figure 4-20). These personalized metabolic predictions also demonstrate the significant amount of heterogeneity in redox metabolism between patient tumors. While the majority of initiatives for precision/personalized cancer treatment has focused on mutations and expression differences in signaling pathway proteins, a greater focus on exploiting metabolic differences between patient tumors is currently being made [253]. For example, heterogeneity in glycolytic metabolism is being used to identify diagnostic biomarkers and treatment strategies for both pancreatic cancer and acute myeloid leukemia patients [254, 255]. Personalized nutrition may also be a viable strategy for manipulating tumor metabolism, including redox metabolism, in individual patients; most NADPH-generation pathways are supplied by glucose and glutamine intake, and folic acid levels may impact folate-dependent reactions [256, 257]. Continued development of genome-scale metabolic models of individual patient tumors will undoubtedly aid in the identification of novel metabolic targets or optimal nutrition strategies for individual cancer patients.

Some differences in predictions arose between these FBA-based metabolic models and previous computational models of redox metabolism, mainly in the relative flux magnitudes of H_2O_2-clearing enzymes. In their ODE-based models of H_2O_2 elimination in Jurkat T cells, Adimora et al. predicted that catalase was insignificant in its contribution to steady-state H_2O_2 clearance; instead, the major contributors to H_2O_2 elimination were predicted to be glutathione peroxidase, peroxiredoxin, and thioredoxin (involving reduction of the protein dithiol pool) [237]. However, the authors only included

peroxisomal catalase in their model construction, and mentioned that the low peroxisomal content of Jurkat cells as well as low diffusion rate of H_2O_2 from the cytosol into peroxisomes could lead to the small predicted contribution of catalase. There are also known differences in enzyme utilization for H_2O_2 clearance based on the observed H_2O_2 concentration; while other enzymes may dominate at lower concentrations, catalase has been shown to be the primary scavenging enzyme at higher H_2O_2 concentrations which can be induced by ionizing radiation therapy and cause double-stranded DNA breaks [258-261]. Prior studies on catalase have focused on the predominant peroxisomal form, as well as its cytosolic form which varies between cell types due to differences in membrane localization [74, 262-264]. Nonetheless, mitochondrial catalase has been detected and shown to mitigate cellular damage caused by mitochondrially-produced ROS [265-267]. Our FBA models predicted that while peroxisomal catalase did not show significant flux differences between radiation-sensitive and radiation-resistant tumors, mitochondrial catalase showed significantly higher fluxes in radiation-resistant tumors. Even though the magnitudes of mitochondrial catalase flux predicted from these FBA models may be higher than expected due to the use of the COMPARTMENTS database to calculate relative catalase distribution as well as the lack of transport constraints on H_2O_2 diffusion between cellular compartments, the observation of higher fluxes in radiation-resistant tumor models is consistent with greater mitochondrial NADPH generation which was also observed.

Despite the recent increase in experimental metabolomics studies towards identifying altered metabolic phenotypes in cancer, an integrated assessment of the 13,000+ human metabolic reactions and their collective impact on individualized treatment response is currently infeasible using solely experimental approaches [268]. Instead, computational

approaches which integrate multi-omic measurements and global reconstructions of human metabolism into predictive models provide tremendous utility in improving our understanding of pathophysiological processes and discovering personalized metabolic targets [10]. In this study, we have shown that flux balance analysis models of individual patient tumors can identify important differences in redox metabolism between radiation-sensitive and -resistant tumors. Specifically, radiation-resistant tumors can upregulate metabolic flux through mitochondrial NADPH-generating and H_2O_2-clearing reactions to increase cellular antioxidant stores, decrease levels of reactive oxygen species, and resist the damaging effects of ionizing radiation treatment. Additionally, these metabolic models identified novel targets in the redox metabolic network that significantly impact antioxidant production and ROS response, both across the entire radiation-resistant cohort as well as in individual patient tumors. These targets represent potential biomarkers for the *a priori* prediction of radiation sensitivity, as well as therapeutic strategies for sensitizing tumors to radiation therapy. Ultimately, the development of personalized metabolic models has the potential to facilitate the clinical management of cancer patients and improve long-term outcomes.

CHAPTER 5. MACHINE LEARNING CLASSIFIERS FOR PREDICTION OF RADIATION RESPONSE

5.1 Introduction

Tumor resistance to radiation therapy has been attributed to multiple pathophysiological mechanisms, including alterations in cellular metabolism, signaling, and DNA damage repair [15, 16, 269]. Despite recent advancements in our biological understanding of radiation resistance, methodologies for the *a priori* prediction of radiation response in individual cancer patients have yet to be implemented clinically. The ability of predicting therapeutic response would ensure that cancer patients only receive optimal therapies tailored to that individual, thereby improving patient outcomes, preventing unnecessary side effects, and decreasing overall healthcare costs.

Statistical methods for identifying clinical biomarkers for predicting radiation therapy response have been used for decades. Early studies focused on correlations between radiation response and clinical factors including tumor staging, primary site, and Karnofsky performance score [31, 33]. The inclusion of physiological parameters such as tumor oxygenation status resulted in improved predictions of radiation sensitivity [35, 36]. Predictors using clinical information later switched from classical statistical methods to newer machine learning algorithms which enabled improvements in accuracy and biomarker detection [270-272]. As methods for transcriptomic analysis improved with microarrays and later RNA-seq, gene expression-based classifiers for radiation response rose to popularity and have recently been curated by the RadiationGeneSigDB database [46, 273-283]. These classifiers span a wide range of feature count (from a handful to

many-thousand), statistical or machine learning method used for biomarker detection, and original gene expression dataset. However, there is currently a lack of radiation response classifiers which integrate multiple -omics modalities. While genomic and transcriptomic data is becoming more widely available for large numbers of patient tumors through initiatives such as The Cancer Genome Atlas, metabolomic data from these tumors is lacking, limiting its use in predictive models for radiation therapy response [2].

Genome-scale metabolic modeling approaches such as flux balance analysis (FBA) are becoming increasingly popular for predicting metabolic phenotypes in individual cancer cell lines or tumors of interest [9, 10, 284]. By combining a curated reconstruction of the human metabolic network with constraints on metabolic reaction activities and an objective function to maximize a particular metabolic phenotype, predictions of steady-state reaction fluxes or metabolite production rates under physiological constraints can be obtained at a genome scale. We previously developed a novel bioinformatics pipeline for integrating genomic, transcriptomic, kinetic, and thermodynamic parameters into personalized FBA models of 716 radiation-sensitive and 199 radiation-resistant patient tumors across multiple cancer types from TCGA. Using these personalized metabolic models, we discovered that radiation-resistant tumors could upregulate levels of reduced redox cofactors such as NADPH and glutathione to increase antioxidant stores and clearance of reactive oxygen species (ROS) generated from ionizing radiation. We then identified novel targets of redox metabolism which could inhibit antioxidant production and ROS-clearance in radiation-resistant cancers; these model predictions were validated by performing siRNA gene knockdown experiments in a panel of matched radiation-sensitive and radiation-resistant cancer cell lines. With these models, we also demonstrated

the ability to predict metabolic phenotypes in individual patient tumors, which could be used to inform personalized diagnostic and therapeutic strategies. Thus, these models provide accurate genome-scale metabolic predictions in tumor samples with little available experimentally-measured metabolomics information.

While machine learning methods have been previously combined with genome-scale metabolic models to improve prediction of metabolic phenotypes, most studies combining these two methodologies have focused on microbiological applications rather than applications to cancer metabolism or predicting treatment outcomes [285]. We hypothesize that output predictions from personalized genome-scale metabolic models of individual patient tumors would not only provide an additional -omics dataset to better understand pathophysiological differences between radiation-sensitive and -resistant tumors at a systems level, but could also be integrated with clinical, genomic, and transcriptomic data to develop accurate machine learning classifiers for the *a priori* diagnosis of radiation sensitivity in cancer patients. To this end, we developed a machine learning pipeline for integrating multi-omic data into machine learning classifiers which outperform previous predictors of radiation sensitivity. Additionally, we utilized recently-developed FBA models of TCGA patient tumors to predict production rates of metabolites across the human metabolic network, and incorporated these into machine learning classifiers to obtain novel metabolite biomarkers associated with sensitivity or resistance to radiation therapy. Finally, we demonstrate the capability to predict optimal multi-omic and non-invasive biomarker panels with maximum diagnostic utility for individual cancer patients.

5.2 Methods

5.2.1 Computational methods

5.2.1.1 TCGA data retrieval and processing

Clinical data from TCGA patients was obtained from the GDC data portal (clinical drug, clinical patient, and clinical radiation files) and the Synapse TCGA_Pancancer project (biological sample files) [2, 286]. Drug names were standardized according to the standard available from the Gene-Drug Interactions for Survival in Cancer (GDISC) database [287]. Categorical clinical features were one-hot encoded before inputting into machine learning classifiers. RNA-Seq gene expression data was obtained from Rahman et al.'s alternative preprocessing method (Gene Expression Omnibus GSE62944) [211]. Data from this preprocessing method showed fewer missing values, more consistent expression between replicates, and improved prediction of biological pathway activity compared to the original TCGA pipeline. Mutation data using the MuTect variant caller was obtained from the GDC data portal [2, 288]. For all data types, only features with at least two unique non-missing values were included.

5.2.1.2 Radiation sensitivity

TCGA samples were classified into radiation-sensitive and radiation-resistant groups according to their reported sensitivity to radiation therapy based upon the RECIST classification method. Patients with a complete or partial response to radiation (greater than 30% decrease in tumor size) were classified as radiation-sensitive, and patients with stable or progressive disease (either less than 30% decrease in tumor size, or increase in tumor size) were classified and radiation-resistant. If a patient received multiple courses of radiation therapy, they were classified based on the response to their first course.

5.2.1.3 Data splitting

The collection of 716 radiation-sensitive and 199 radiation-resistant samples was randomly split into training+validation (80%) and testing (20%) groups. Within the training+validation group, 5-fold cross validation was performed to identify optimal hyperparameter values. The training (80%) group was used for training the model with a given set of hyperparameters; within this training group, 87.5% was directly used for training, and 12.5% was used to identify the optimal iteration to perform early stopping during training. The validation (20%) group was used to assess model performance with the given set of hyperparameters. The averaged validation performance across all 5 folds was used to determine the optimal set of hyperparameters; once this optimal set was determined, the model was retrained on the training+validation group, and the testing group was used to assess overall model performance. 20 iterations of randomized training+validation/testing splitting were performed to analyze model predictions and performance metrics over multiple instances. All data splits were performed using stratified shuffle splitting, where the proportion of radiation-sensitive and -resistant samples were kept the same between the two groups.

5.2.1.4 Gradient boosting classifiers

XGBoost (v0.90) was used to develop gradient boosting machine (GBM) base and meta learner classifiers [289]. N_d base classifiers were trained using individual -omic data types (e.g. clinical, gene expression, mutation, or FBA-predicted metabolomics), where N_d is the number of individual data types. For individual samples, each of the N_d base classifiers outputs the predicted probability of radiation resistance (p_1, p_2, …, p_{N_d}) using

features from the individual data type. Each base classifier receives the same training/validation/testing split of samples.

Subsequently, for each sample within the 5 validation sets used with the N_d base learners, the best individual base learner is determined using the sample's known radiation class and predicted probabilities of radiation resistance from the N_d base learners (e.g. the best base learner for a radiation-resistant sample is that which outputs the highest probability of radiation resistance for that sample). A meta learner is then trained using these samples to predict the best base learner based on sample features from all N_d data types. The meta learner outputs the predicted probability that each of the N_d base learners is the best for a particular sample (w_1, w_2, …, w_{Nd}). This multi-class classification is expedited by only using features from the N_d data types with non-zero SHAP values from the N_d individual base learners (see explanation of SHAP values below). Because validation samples from the 5-fold cross validation were never directly used in training the base learners (only used in assessing the accuracy of base learners and determining optimal hyperparameter values), they can be used to train this meta learner without overfitting or inflation of model performance metrics. Note that, once the meta learner is trained using the predicted probabilities from the base learners, the base learners and meta learner act independently of each other when used on new testing samples.

Each testing sample is run through 1) all N_d base learners to obtain the predicted probabilities of radiation resistance using each of the N_d individual data types (p_1, p_2, …, p_{Nd}); and 2) the meta learner to obtain the predicted probabilities that each of the N_d base classifiers is the best classifier for that sample (w_1, w_2, …, w_{Nd}). To obtain the final predicted probability of radiation resistance for the testing sample, the weighted average

of the base learner probabilities is taken, with the meta learner probabilities as weights: $p = w_1p_1 + w_2p_2 + \ldots + w_{Nd}p_{Nd}$. If this final probability is greater than 0.5, the testing sample is classified as radiation-resistant; otherwise, it is classified as radiation-sensitive.

5.2.1.5 Bayesian optimization

To determine the optimal combination of GBM hyperparameters for both the base learner and meta learner classifiers, Bayesian optimization was used. This iterative approach automates the search for optimal hyperparameter values by calculating an acquisition function which provides the expected benefit of sampling a particular point in hyperparameter space on the overall search for optimal hyperparameters with minimal cross-validation error. At each iteration, the point in hyperparameter space with the largest acquisition function value is chosen, 5-fold cross validation is used to determine the performance of those particular hyperparameters, and the acquisition function is updated to then determine which next point in hyperparameter space will be sampled. Hyperopt (v0.1.2) was used to perform Bayesian optimization [290]. Table 5-1 provides the 8 GBM hyperparameters chosen for optimization of both base learner and meta learner classifiers, with the ranges of values in the hyperparameter search space. 2_8=256 iterations of Bayesian optimization were performed for each classifier.

Table 5-1. Hyperparameter ranges for Bayesian optimization with gradient boosting classifiers.

Parameter Name	Distribution	Lower Bound	Upper Bound
eta	log-uniform	0.01	0.5
gamma	log-uniform	0	5
max_depth	uniform	1	11
subsample	uniform	0.5	1
colsample_bytree	uniform	0.5	1
colsample_bylevel	uniform	0.5	1
reg_lambda	log-uniform	1	4
reg_alpha	log-uniform	0	1

5.2.1.6 Model performance metrics

Weighted log loss was used as the performance metric to be minimized while performing 5-fold cross validation and Bayesian optimization to determine optimal hyperparameter values for each base learner:

$$\text{Weighted Log Loss} = \frac{1}{N_S}\sum_{i=1}^{N_S}\left[-\left(w_R y_i \log(p_i) + (1-y_i)\log(1-p_i)\right)\right] \quad (27)$$

where y_i is the true class label of sample i (y_i=0 if sensitive, y_i=1 if resistant), p_i is the predicted probability of sample i being radiation resistant (belonging to class 1), w_R is the weight given to radiation-resistant samples (w_R = # sensitive samples / # resistant samples), and N_S is the total number of samples. The weight given to radiation-resistant samples accommodates for the fact that there are more radiation-sensitive samples than radiation-resistant samples, and prevents classifiers from focusing on optimizing performance exclusively on radiation-sensitive samples rather than all samples equally.

Multiclass log loss was used as the performance metric for the meta learner:

$$\text{Multiclass Log Loss} = \frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{k=0}^{N_k} y_{i,k}\log\left(p_{i,k}\right) \tag{28}$$

where $y_{i,k}$ is 1 if dataset k is the true optimal dataset of sample i and 0 otherwise, $p_{i,k}$ is the predicted probability of dataset k being the optimal dataset of sample i, N_S is the total number of samples, and N_k is the total number of datasets.

Final model performance was assessed on testing samples across the 20 iterations of randomized training+validation/testing splitting. The following performance metrics were used:

1. Weighted log loss (explained above)
2. Area under the receiver operating characteristic (ROC; plot of true positive rate vs. false positive rate) curve (AUROC)
3. Balanced accuracy, an accuracy metric that corrects for unequal numbers of radiation-sensitive and -resistant patients:

$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{TP}{TP+FN}+\frac{TN}{TN+FP}\right) \tag{29}$$

4. Sensitivity:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{30}$$

5. Specificity:

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{31}$$

6. Positive predictive value:

$$\text{Positive Predictive Value} = \frac{TP}{TP + FP} \tag{32}$$

7. Negative predictive value:

$$\text{Negative Predictive Value} = \frac{TN}{TN + FN} \tag{33}$$

5.2.1.7 SHAP values

The importance of individual features towards the prediction of radiation sensitivity, both averaged across all samples as well as for individual samples, was determined by calculating SHAP values for each classifier. Each SHAP value represents the change in the predicted probability of radiation resistance for patient *i* attributed to feature *j* [291]. Features with positive SHAP values for patient *i* signify those where the particular value of feature *j* attributed to patient *i* is such that it increases patient *i*'s predicted probability of radiation resistance; larger absolute SHAP values indicate features with larger contributions to the predicted probability (either negatively or positively). Mean absolute SHAP values across all samples provide an indication of the overall importance of a particular feature in the classifier's prediction of radiation sensitivity. SHAP values were averaged across 20 training+validation/testing splits by a weighted average, with weights proportional to the inverse of the weighted log loss performance metric on the testing set for that split. Values were normalized by the difference between prior and posterior probabilities of radiation resistance for each sample. SHAP v0.29.1 was used to calculate SHAP values [292].

5.2.1.8 Comparison of machine learning algorithms

sklearn.ensemble.RandomForestClassifier() and sklearn.linear_model.LogisticRegression() functions with scikit-learn v0.21.2 were used to implement random forest and logistic regression with L1 regularization classifiers, respectively. Keras v2.3.1 was used to implement the neural network with L1 regularization classifier. Weighted log loss was used as the loss function for the neural network classifier, and early stopping was performed in the same manner as for the GBM classifier. Missing values were imputed and scaled using sklearn.impute.SimpleImputer() and sklearn.preprocessing.StandardScaler() functions, respectively, before training with the random forest, logistic regression, and neural network algorithms. Table 5-2, Table 5-3, and Table 5-4 provide the hyperparameters and value ranges used for Bayesian optimization with each algorithm. 20 training+validation/testing splits were performed for each algorithm except the neural network, where during 7/20 splits some samples were explicability dropped during training/testing, preventing the performance from these splits to be compared with other algorithms.

Table 5-2. Hyperparameter ranges for Bayesian optimization with the random forest classifier.

Parameter Name	Distribution	Lower Bound	Upper Bound
n_estimators	log-uniform	1	1000
criterion	choice	entropy, gini	
max_depth	uniform	1	11
min_samples_split	uniform	0	1
min_samples_leaf	uniform	0	0.5
max_features	choice	log2, sqrt	

Table 5-3. Hyperparameter ranges for Bayesian optimization with the logistic regression classifier with L1 regularization.

Parameter Name	Distribution	Lower Bound	Upper Bound
C	log-uniform	0.001	1000

Table 5-4. Hyperparameter ranges for Bayesian optimization with the neural network classifier with L1 regularization.

Parameter Name	Distribution	Lower Bound	Upper Bound
Number of layers	uniform	1	5
Neurons per layer	choice	8, 16, 32, 64, 128, 256, 512, 1024	
activation_function	choice	elu, relu, sigmoid	
optimizer	Choice	adam, rmsprop, sgd	
l1	log-uniform	0.000001	0.1
dropout	uniform	0	0.5

5.2.1.9 Comparison of gene expression datasets

11 oxic radiation response gene expression signatures from RadiationGeneSigDB were compared to the top genes from the gene expression classifier (782 genes with cumulative SHAP value sum of 95%). Gene names from RadiationGeneSigDB signatures were converted to Entrez gene ID's and gene symbols. Those genes where a matching Entrez gene ID or gene symbol could not be found were removed. Additionally, those genes that were not in both TCGA gene expression and CCLE gene expression datasets were removed.

To compare performance of gene expression signatures on radiation sensitivity prediction using TCGA data, GBM classification models predicting radiation-sensitive vs. radiation-resistant classes of TCGA samples were trained using expression data from the subset of genes for each signature. The performance of these models was assessed via both weighted log loss and AUROC. To compare performance of gene expression signatures on radiation sensitivity prediction using CCLE data, GBM regression models predicting radiation response (reported as area under the curve of survival vs. radiation dose; [293]) of CCLE cell lines were trained using expression data from the subset of genes for each signature. The performance of these models was assessed via both mean absolute error (MAE) and mean squared error (MSE).

5.2.1.10 Flux Balance Analysis (FBA)

Generation of personalized FBA models of individual TCGA samples was performed as described above. To predict the maximum production of particular metabolites in TCGA models, the following objective function was used:

$$1\ \mathrm{met[all]} \rightarrow \tag{34}$$

where "met" is the metabolite to be maximized, and "[all]" represents the maximization of the objective function across all cellular compartments where the metabolite is located. The modeled external compartment contained all metabolites found in DMEM/F-12 cell culture media (Thermo Fisher #11320) as well as fetal bovine serum (FBS) to match the cell culture media used for experimental validation [219].

5.2.1.11 Code availability

Code for the generation of machine learning classifiers for radiation response using multi-omic TCGA data is available at https://github.com/kemplab/ML-radiation. Jupyter notebooks are available for 1) training/testing of base learner classifiers; 2) training/testing of meta learner classifiers; 3) analysis of classifier predictions and feature importances.

5.2.2 Experimental methods

5.2.2.1 Cell culture

Table 5-5 provides the matched radiation-sensitive and radiation-resistant cell lines used for experimental validation of metabolites levels predicted from FBA. All cell lines were maintained in DMEM/F-12 cell culture media (Thermo Fisher #11320) with 10% fetal bovine serum (Sigma-Aldrich #F4135) at 37°C and 5% CO_2, and were free of *Mycoplasma*.

Table 5-5. Matched radiation-sensitive and radiation-resistant cell lines.

Cancer Type	Radiation-Sensitive	Radiation-Resistant	Notes	Source
Breast (BRCA)	MDA-MB-231 NQO1(-)	MDA-MB-231 NQO1(+)	Stable NQO1 expression was restored in NQO1(-) cells to create NQO1(+) cells.	Dr. David Boothman, Indiana University[220]
Colon (COAD)	SW620	SW480	Primary tumor (SW480) and lymph node metastasis (SW620) from the same patient, with different radiation sensitivities.	ATCC
Glioblastoma (GBM)	M059J	M059K	Both isolated from same tumor specimen. M059J cells lack DNA-PK activity, rendering them more radiation-sensitive.	ATCC
Head and Neck (HNSC)	SCC-61	rSCC-61	rSCC-61 cells were derived from SCC-61 cells after repeated radiation exposure and selection of surviving colonies.	Dr. Cristina Furdui, Wake Forest University[175]

5.2.2.2 Metabolomics

Three biological replicates of each cell line were grown in separate T-25 flasks with the cell culture conditions described above. Cell pellets with approximately 1 million cells were obtained from trypsinization, centrifugation, and removal of supernatant. Samples were reconstituted in 90% MeOH, 10% H_2O at a ratio of 200 μL/1 million cells. Aliquots of the supernatant were combined to create a pooled sample used for quality control. Aliquots of the samples were transferred to LC vials and stored at 4°C.

Hydrophilic interaction chromatography-tandem mass spectrometry (HILIC-MS/MS) untargeted metabolomics was performed. Chromatography parameters were as follows: BEH HILIC Column, 150 mm X 2.1 mm, 1.7 μm; mobile phase A: 80% H_2O / 20% ACN, 10mM ammonium formate, 0.1% FA; mobile phase B: 100% ACN, 0.1% FA; column temperature: 40°C; 2 μL sample injection. MS parameters were as follows: resolution: 240,000; scan range: 70-1050 m/z; polarity: positive/negative; AGC target: 1e5. MS_2 parameters were as follows: isolation window: 0.8 m/z; detector: Orbitrap; polarity: positive/negative; fragmentation method: HCD; collision energy: 15, 30, 45; resolution: 30,000.

Compound Discoverer 3.1 was used to perform quality control, metabolite identification, and quantification of metabolite levels. Results for positive and negative ion modes were combined. Metabolites with no identified name were initially removed from the analysis. If duplicate metabolites with the same identification were obtained, then the entry with the largest maximum area was used. KEGG ID's for each metabolite were manually identified based on metabolite name, molar mass, and chemical formula. Metabolites from experimental metabolomics were matched to those from FBA analysis by matching KEGG ID's.

For the comparison of model-predicted and experimentally-measured metabolite values, all metabolites within the following Recon3D subsystems that were matched with experimental metabolites were included in the analysis:

- Nucleotide Metabolism: "Nucleotide interconversion", "Nucleotide salvage pathway", "Pentose phosphate pathway", "Purine catabolism", "Purine synthesis", "Pyrimidine catabolism", "Pyrimidine synthesis"

- Lipid Metabolism: “Cholesterol metabolism”, “Fatty acid oxidation”, “Fatty acid synthesis”, “Glycosphingolipid metabolism”, “Phosphatidylinositol phosphate metabolism”, “Sphingolipid metabolism”, “Steroid metabolism”
- Cysteine/Antioxidant Metabolism: “Glutathione metabolism”, “ROS detoxification”, plus metabolite “Lipoamide”
- Immune System Mediators: “Arachidonic acid metabolism”, “Eicosanoid metabolism”

5.3 Results

5.3.1 Dataset-independent ensemble for radiation response classification

716 radiation-sensitive and 199 radiation-resistant TCGA patient tumors were available for classification of radiation sensitivity (Figure 5-1). These samples are divided into training+validation and testing sets, where the testing set is used solely to evaluate model performance. 5-fold cross validation using training and validation sets is used to determine optimal hyperparameter values. Stratified shuffling is performed at each data split to ensure equal distribution of radiation-sensitive and -resistant samples at each step. 20 randomized iterations of training+validation/testing splits are performed to obtain unbiased averages of model performance metrics across multiple data splits. When multiple datasets are integrated for prediction of radiation response, optimal classifier performance was found when implementing a dataset-independent ensemble architecture, where each dataset is first trained on an independent “base learner” machine learning classifier (Figure 5-2). Afterwards, outputs of each base learner are integrated into a “meta learner” that determines which datasets provide the most predictive value for each

individual sample. Finally, the base learner and meta learner predictions are used to determine the ultimate classification of each sample as either a radiation-sensitive or radiation-resistant tumor.

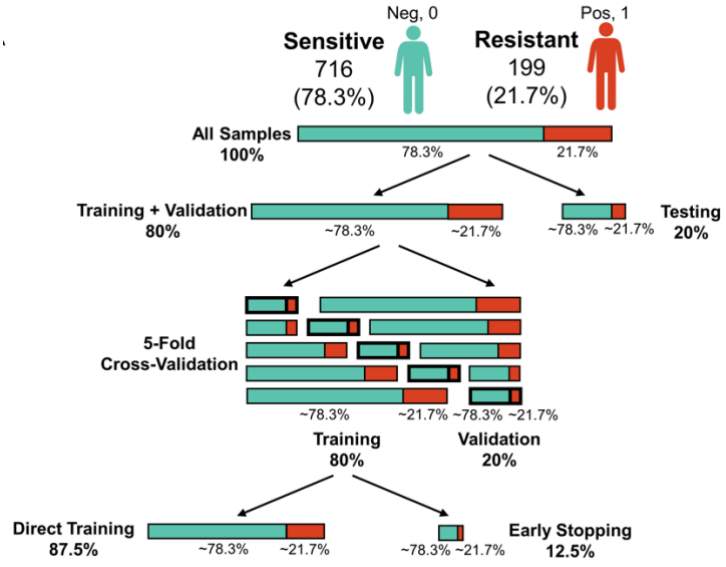


Figure 5-1. Splitting of data into training (both direct training and early stopping), validation, and testing sets.

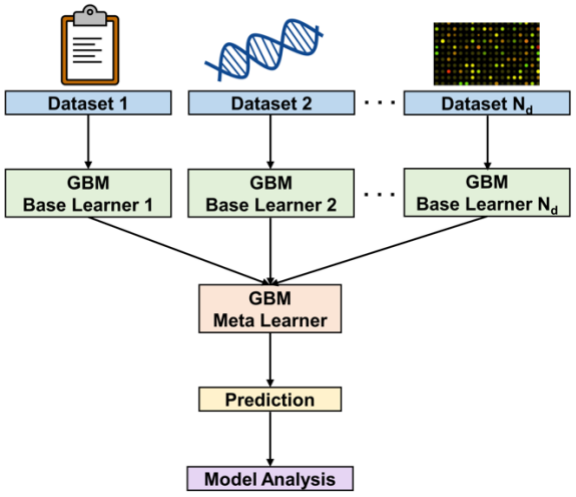


Figure 5-2. Dataset-independent ensemble architecture, with base learners for each dataset and one meta learner for integration of base learner outputs.

Each base learner is composed of a gradient boosting machine (GBM) that performs two-class classification (either radiation-sensitive or -resistant) using features from an individual dataset (Figure 5-3). Besides providing optimal performance accuracy, GBM classifiers using decision tree ensembles have many useful characteristics compared to other machine learning algorithms, including embedded feature selection, capability of handling missing values (which is common in clinical datasets), and efficient management of high-dimensional datasets (where the number of features greatly exceeds the number of samples) [289, 294, 295]. Bayesian optimization is performed to determine the optimal hyperparameter values for the GBM classifier; by calculating an acquisition function that determines the most informative point in hyperparameter space to sample at each iteration, this approach yields better performance compared to traditional approaches such as grid search or random search [290, 296]. At each iteration of Bayesian optimization, 5-fold cross validation is used to calculate the performance of the set of hyperparameters, measured by the mean weighted log loss plus one standard error; greater weight is given to

radiation-resistant samples in the weighted log loss to ensure that the classifier performances equally well on both classes despite the lower number of radiation-resistant samples. During model training, early stopping is employed to prevent overfitting. After the classifier has been trained, it is run on the testing set using the optimal set of hyperparameters.

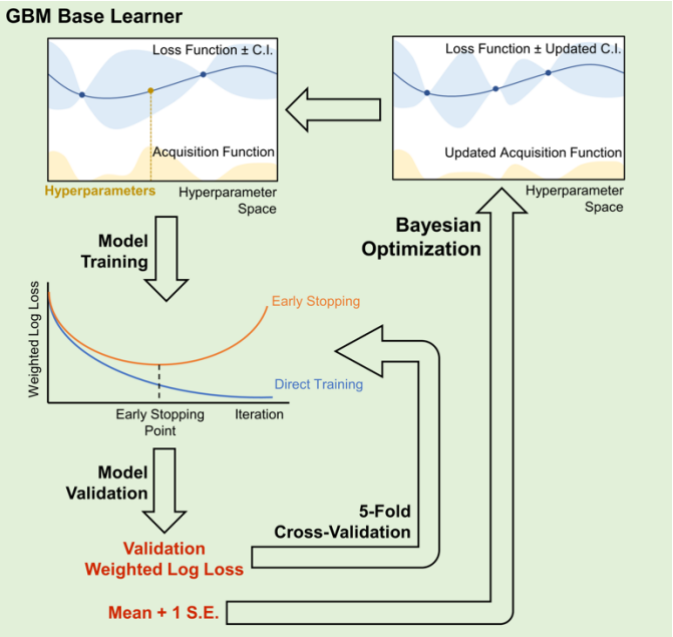


Figure 5-3. GBM base learner performing two-class classification of radiation sensitivity or resistance. For each iteration of Bayesian optimization and set of hyperparameters, 5-fold cross validation is performed. The set of hyperparameters with the lowest mean weighted log loss plus 1 standard deviation is chosen. Early stopping is employed during training to prevent model overfitting.

By comparing each base learner's prediction to the true class of validation set samples, the meta learner is trained to determine which base learner/dataset provides the most accurate prediction of radiation sensitivity for each new sample (Figure 5-4). This meta learner performs an N_d-class classification, where N_d is the number of independent

base learners/datasets, and the output is N_d probabilities that each base learner/dataset provides the most accurate prediction (all N_d probabilities sum to 1). The features this meta learner is trained on include all features from the N_d datasets which have non-negative SHAP values from their respective base learners; features which do not impact the base learner predictions are not included, which increases the training speed while maintaining accuracy. The implementation of this GBM-based meta learner is otherwise similar to that of each base learner, utilizing Bayesian optimization and 5-fold cross validation to determine optimal hyperparameter values. After training of base and meta learners, each testing set sample is run through all N_d base learners and the meta learner (Figure 5-5). Each of the base learners calculates a probability p_i (where $i \in (1,N_d)$) of radiation resistance ($0 \leq p_i \leq 1$), and the meta learner calculates the weights w_i for each p_i based on how accurate it predicts each base learner to be on this sample. The final probability of radiation resistance assigned to the sample is the weighted average $p = w_1p_1 + w_2p_2 + \ldots + w_{Nd}p_{Nd}$; samples with $p < 0.5$ are classified as radiation-sensitive, while samples with $p > 0.5$ are classified as radiation-resistant. Because the meta learner is utilized independently of base learners when running testing set samples, it can be used to determine which datasets provide the most utility for a given patient, which can inform strategies for obtaining optimal clinical and biological measurements at a personalized level.

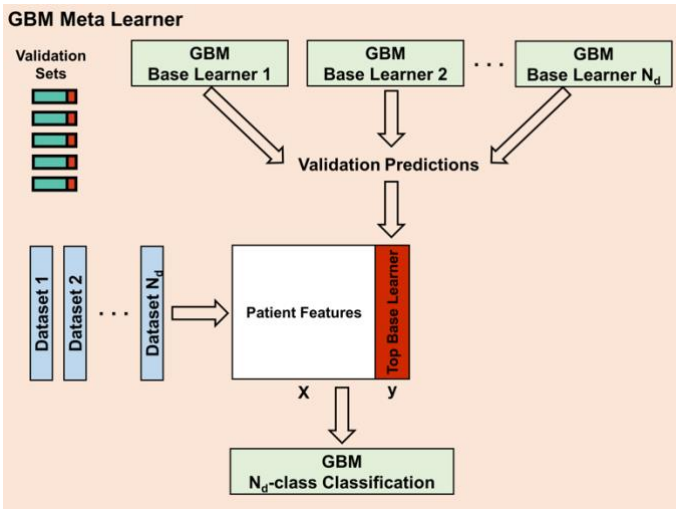


Figure 5-4. GBM meta learner performing N_d-class classification of the most accurate base learner/dataset for each sample, where N_d is the number of independent base learners/datasets. The meta learner is trained using features and base learner predictions on the validation sets.

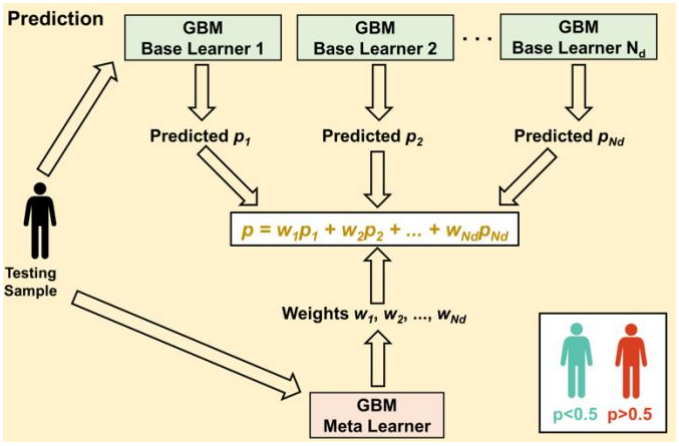


Figure 5-5. Prediction of radiation response for each testing set sample. Base learners provide the predicted probabilities of radiation resistance (p_i), and the meta learner provides the weights for each base learner probability (w_i) based on its predictions of the accuracy of each base learner on the individual sample. The ultimate probability of radiation resistance for the sample is the weighted sum p =

$w_1p_1 + w_2p_2 + \ldots + w_{Nd}p_{Nd}$; samples with $p < 0.5$ are classified as radiation-sensitive, while samples with $p > 0.5$ are classified as radiation-resistant.

To demonstrate the importance of these different components of the machine learning pipeline, the performance of a classifier trained on clinical, gene expression, and mutation data from TCGA was compared with versus without these different components. Figure 5-6 shows that the dataset-independent ensemble architecture performs better than initially combining all three datasets and training on a single GBM classifier; this increase in performance holds across multiple different performance metrics, including weighted log loss, area under the receiver operating characteristic curve (AUROC), balanced accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Performing Bayesian optimization provides an additional increase in performance compared to using default hyperparameter values (Figure 5-7). Furthermore, the use of GBM's as the base and meta learners provides greater performance compared to other commonly used machine learning algorithms with embedded feature selection, including random forests, logistic regression with L1 regularization, and neural networks with L1 regularization (Figure 5-8). Although the performance of this pipeline varies between patients of different cancer types, there was no significant correlation between performance and number of samples per cancer type or proportion of radiation-resistant samples per cancer type (Figure 5-9). Model performance was lowest for samples within the minority class of heavily-skewed cancer types, such as radiation-resistant samples of the largely radiation-sensitive breast cancer (BRCA; Figure 5-10). However, for more balanced cancer types such as lung adenocarcinoma (LUAD), there was no observable difference in performance between radiation-sensitive and -resistant samples. Overall, this

machine learning architecture is a robust platform for integrating multi-omic data and providing accurate predictions of radiation sensitivity in individual patient tumors.

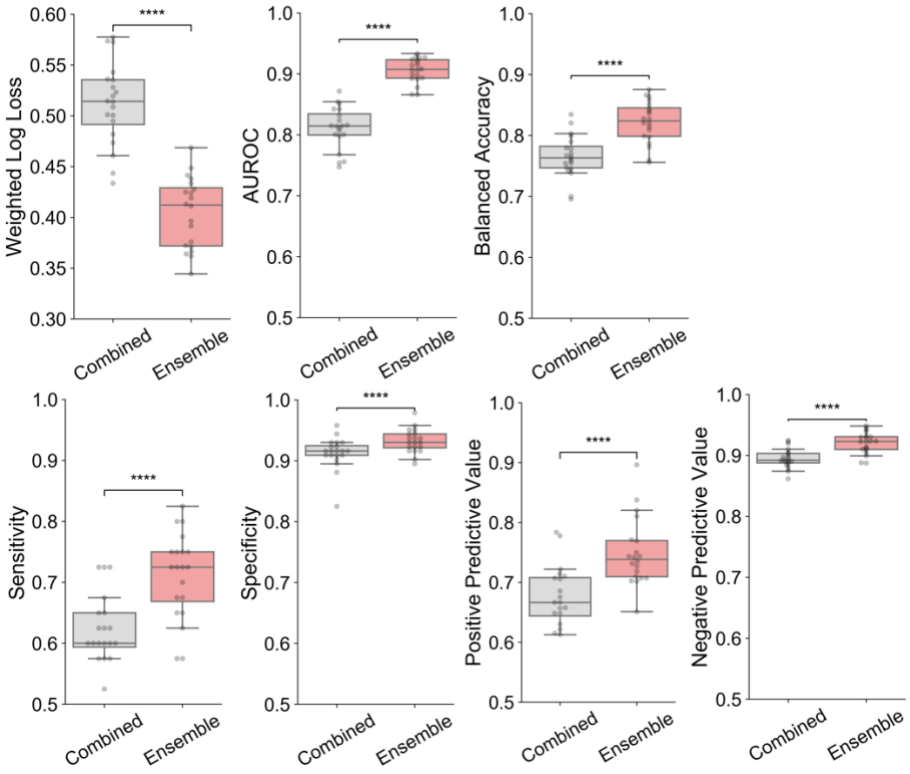


Figure 5-6. Performance comparison of classifiers trained on Clinical + Gene Expression + Mutation data from TCGA, either by combining datasets together before training on a single classifier, or by training on the dataset-independent ensemble architecture.

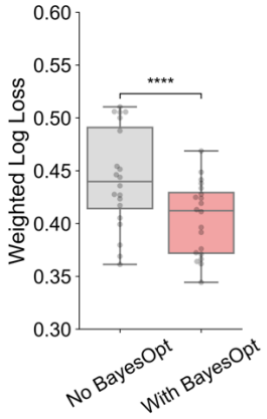


Figure 5-7. Performance comparison of classifiers trained on Clinical + Gene Expression + Mutation data from TCGA, either without (using default hyperparameters) or with Bayesian optimization.

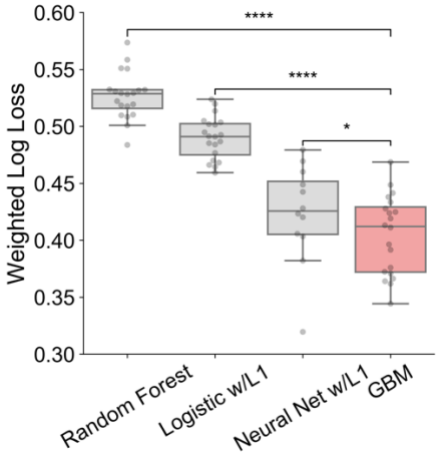


Figure 5-8. Performance comparison of classifiers trained on Clinical + Gene Expression + Mutation data from TCGA using different machine learning algorithms.

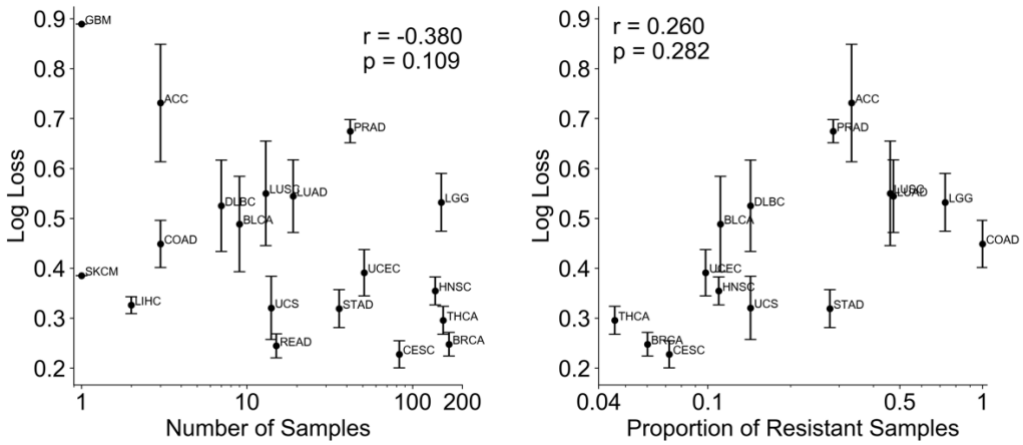


Figure 5-9. (Left) Correlation between sample log loss and number of samples within each cancer type. (Right) Correlation between sample log loss and proportion of radiation-resistant samples within each cancer type.

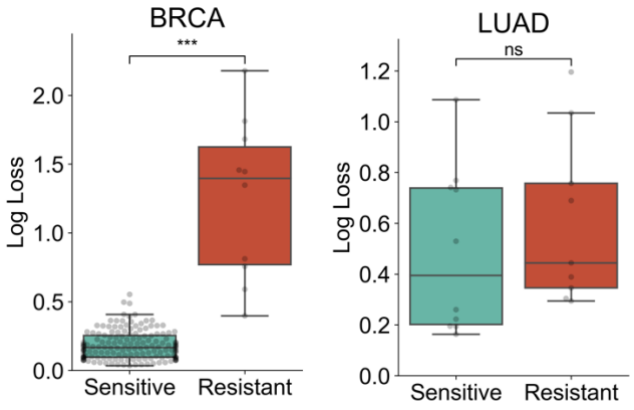


Figure 5-10. Comparison of log loss between radiation-sensitive and radiation-resistant (left) breast cancer (BRCA) samples, and (right) lung adenocarcinoma (LUAD) samples.

5.3.2 Gene expression classifier implicates cellular metabolism

The development of biomarker panels for radiation response has mainly focused on using gene expression data from cancer cell lines or tumors as features [46, 273-283]. To compare our machine learning pipeline with previous classifiers for radiation response and

identify novel gene expression biomarkers, a classifier using RNAseq gene expression data from radiation-sensitive and -resistant TCGA tumors was developed. 782 of the 22,819 genes in the dataset (3.43%) were identified as important in the classification of radiation response, as determined by having a cumulative 95% sum of the classifier's mean absolute SHAP values (mean $|\Delta P|$ across all tumors; Figure 5-11). Among the 50 genes with the largest feature importances, 10 were previously identified as having a role in the tumor response to ionizing radiation therapy [297-306]. To determine whether the identified group of 782 genes has more predictive value than previously identified gene sets in RadiationGeneSigDB, machine learning classifiers were trained using only the genes from each respective gene set, and the predictive accuracy of each classifier was compared (Figure 5-12). When performing a classification task of differentiating radiation-sensitive and -resistant TCGA tumors, our group of 782 genes has the highest accuracy of all gene sets; however, this is likely due to our gene list being discovered by using the TCGA dataset, whereas other gene lists were discovered from other gene expression datasets. To eliminate this potential bias, a separate regression task was performed to predict the area under the curve of survival vs. radiation dose for 524 cancer cell lines from the Broad Institute Cancer Cell Line Encyclopedia (CCLE) [307]. Our group of 782 genes performed among the best gene lists on this separate dataset, and there appeared to be a threshold where gene lists of less than 100 genes could not accurately predict radiation response among the CCLE cell lines.

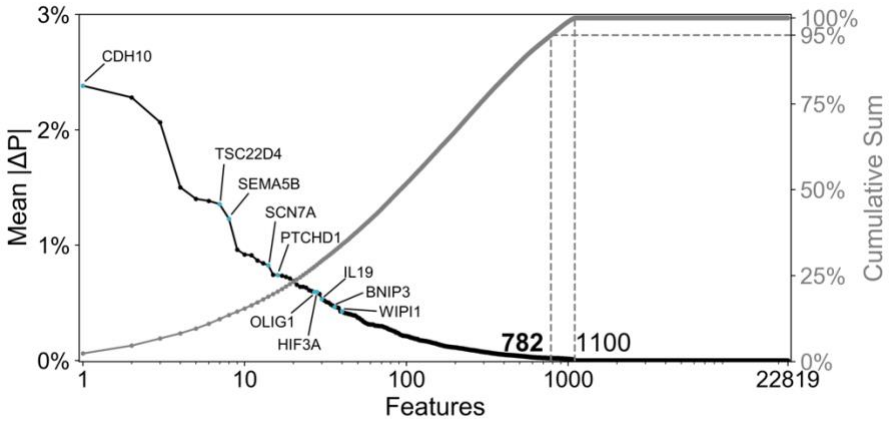


Figure 5-11. (Left; black overlay) Feature importance scores for individual gene expression features. Scores are calculated as the average absolute SHAP value (mean |ΔP|), which is the absolute value of the change in predicted probability of radiation resistance attributed to that feature, averaged across all tumor samples; features with larger scores have greater impacts on the classifier's prediction of radiation sensitivity. Features are ordered by decreasing feature importance score. Among the top 50 features, those with previous literature suggesting a role in tumor radiation response are annotated. (Right; gray overlay) Cumulative sum of feature importance scores. Those features within the 95% cumulative sum of feature importance scores (bolded number) are kept for further analysis.

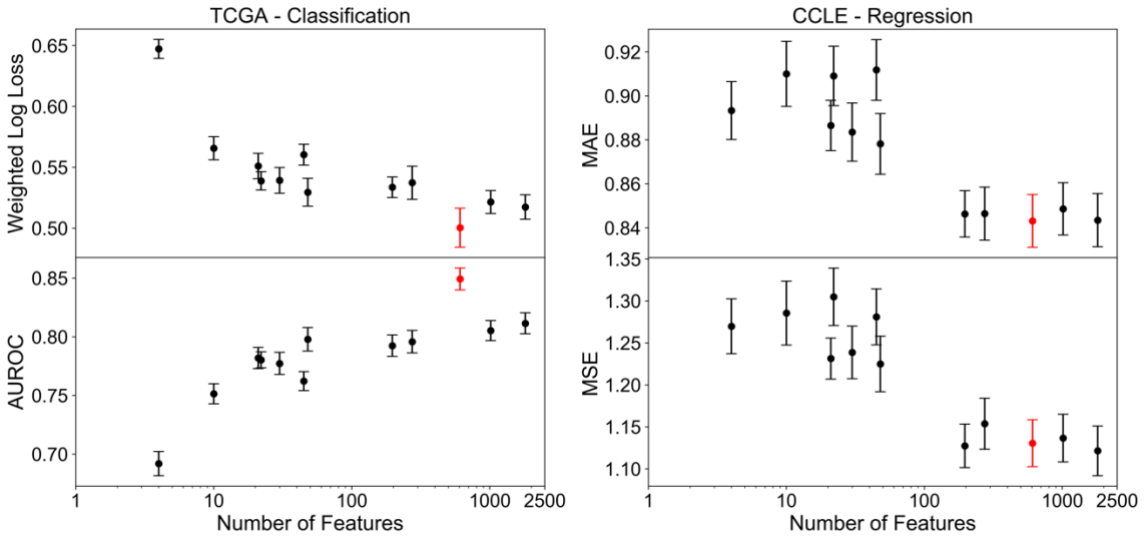


Figure 5-12. Comparison of predictive value of identified gene expression biomarkers (red) to previously identified gene sets in RadiationGeneSigDB (black). (Left) Performance on classification task using gene expression features from TCGA patient tumors, assessing performance by weighted log loss (top) and area under the ROC curve (bottom). (Right) Performance on regression task using gene expression features from CCLE cancer cell lines, assessing performance by mean absolute error (top) and mean squared error (bottom).

To determine how these 782 genes are involved in the biological response to radiation therapy, the enrichment of gene sets representing the Hallmarks of Cancer was determined (Figure 5-13) [308, 309]. The “deregulating cellular energetics” hallmark was the only with significantly greater enrichment than random chance, suggesting that cellular metabolism may play an important role in radiation response. Even though DNA damage is one of the most well-characterized effects of ionizing radiation therapy, the “genome instability & mutation” hallmark had a substantially lower enrichment than random chance [310]. To determine if the positive/negative enrichment of these hallmarks was shared with

other previously-established gene sets for radiation response, hierarchical clustering of the hallmark enrichment rank for each gene set was performed (Figure 5-14). There appeared to be two major clusters of gene sets: a larger cluster having a very high rank for the "genome instability & mutation" hallmark (value close to 1; greater enrichment than chance), and a smaller cluster having a very low rank (value close to 10; lower enrichment than chance). This dichotomy suggests that although the biological response to ionizing radiation therapy certainly involves genomic instability and DNA-damage repair, other biological processes such as cellular metabolism, which have garnered less attention in relation to radiation response, may play critical roles as well [311-313].

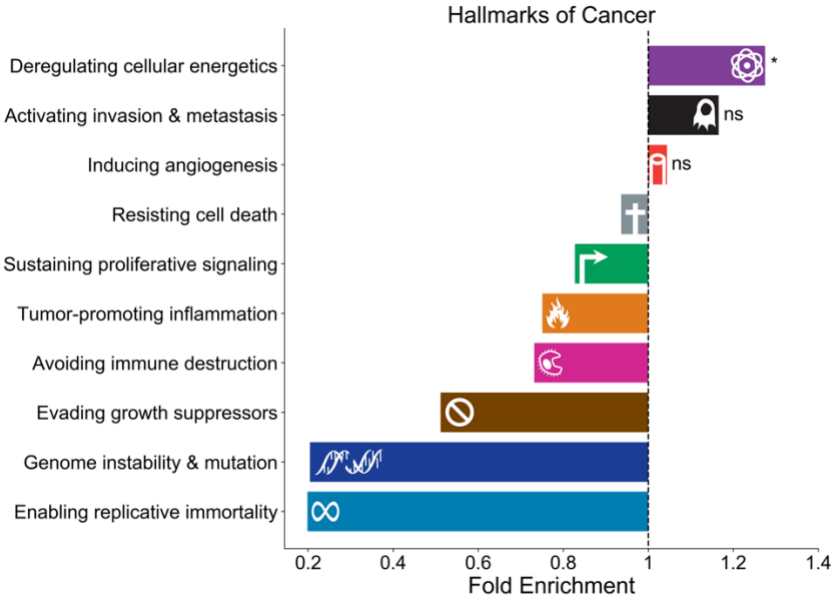


Figure 5-13. Enrichment of hallmarks of cancer by identified gene expression biomarkers. Fold enrichment > 1 indicates hallmarks with greater enrichment than by chance; fold enrichment < 1 indicates hallmarks with lesser enrichment than by chance.

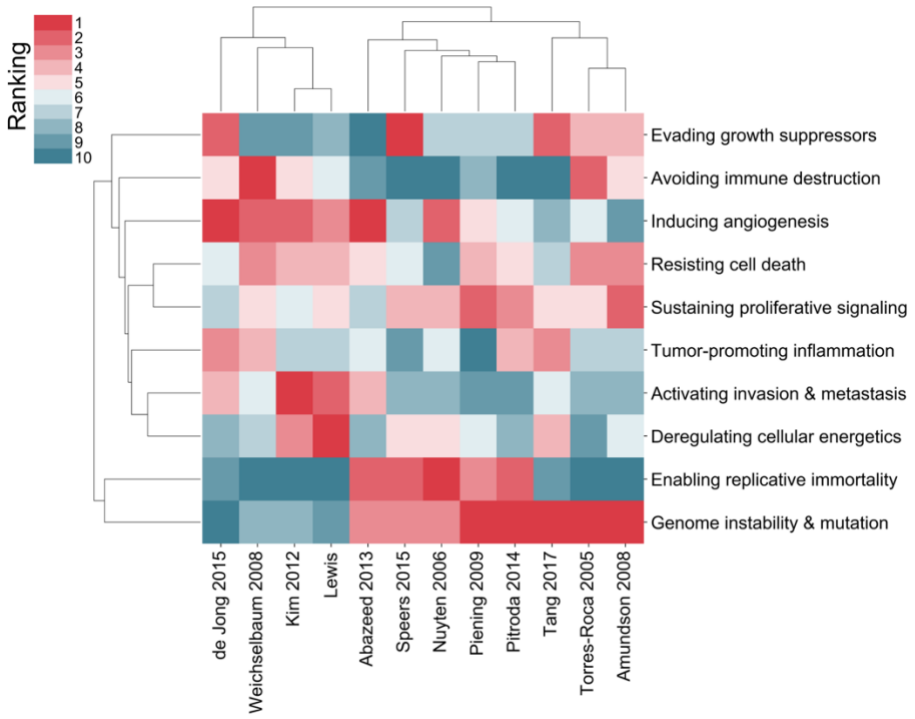


Figure 5-14. Hierarchical clustering of hallmark enrichment ranks based on both gene set (column) and hallmark (row). Enrichment ranks are the rank of each hallmark based on fold enrichment, from high fold enrichments (rank close to 1, greater enrichment than chance) to low fold enrichments (rank close to 10, lower enrichment than chance).

The enrichment of cancer expression modules by our group of 782 genes also suggests an involvement of cellular metabolism in radiation response (Figure 5-15) [314]. Module #354, which involves amino acid and sulfur metabolism, showed a much greater enrichment than all other modules; the sulfur-containing amino acid cysteine is heavily involved in redox metabolism by the presence of its thiol group, and is a precursor for the production of glutathione, an important antioxidant [315, 316]. 42 other expression modules showed significant enrichment and involved disparate metabolic processes

including central carbon metabolism, lipid and fatty acid metabolism, and oxidoreductase activity. To gain better insight into which specific metabolic pathways may show increased or decreased activity in radiation-resistant tumors, the enrichment of individual metabolic subsystems within the Recon3D human metabolic reconstruction was determined (Figure 5-16) [7]. For subsystems with significant enrichment, the SHAP values for each individual gene were analyzed to determine whether increased expression was associated with increased probability of radiation sensitivity or resistance. Several significant metabolic subsystems involved the metabolism of sugars including glucose, fructose, mannose, and galactose, as well as the production of phospholipids and triglycerides. Additionally, the majority of genes involved in these metabolic pathways showed positive correlations between expression and radiation resistance, suggesting that these pathways may be upregulated in radiation-resistant tumors. Together, analysis of this gene expression classifier suggests that radiation-resistant tumors exemplify dysregulation in their cellular metabolic networks, and that additional features involving the cellular metabolism of radiation-sensitive and -resistant tumors will provide significant benefit in developing machine learning classifiers for radiation sensitivity.

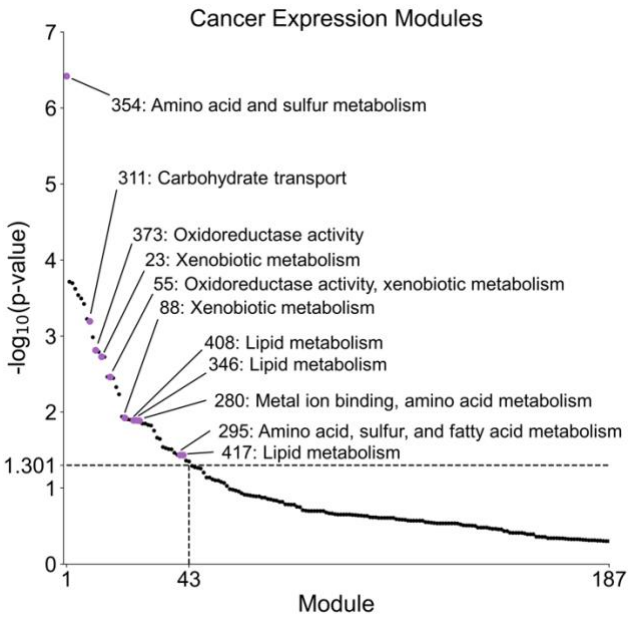


Figure 5-15. Enrichment of cancer expression modules by identified gene expression biomarkers. Modules are ordered by increasing p-value (decreasing statistical significance). Modules relevant to cellular metabolism are annotated with their number and descriptions.

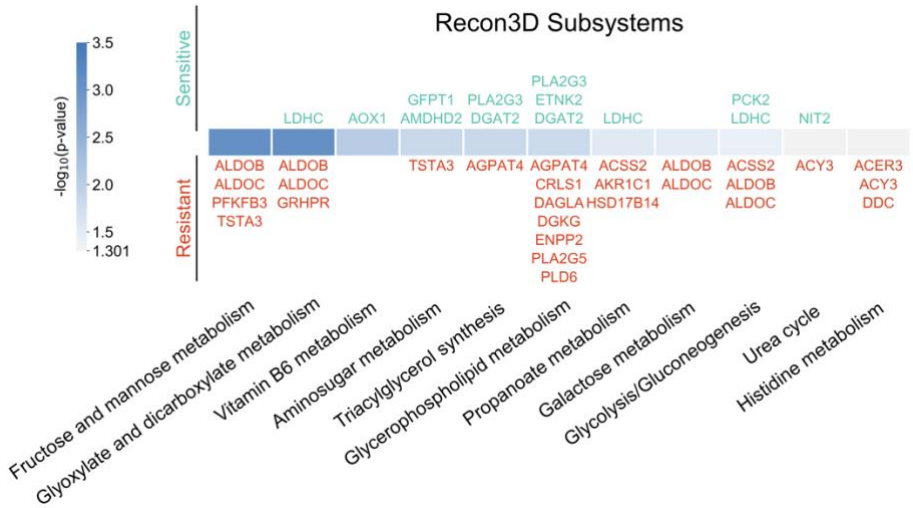


Figure 5-16. Enrichment of Recon3D metabolic subsystems by identified gene expression biomarkers. Subsystems are ordered by increasing p-value (decreasing

statistical significance). Genes within each subsystem are annotated above or below colored significance bars based on whether their expression is positively correlated with predicted sensitivity (green) or resistance (red) to radiation therapy, respectively.

5.3.3 Genome-scale metabolic models accurately predict metabolite production

Although the gene expression and mutations of metabolic genes are directly available from TCGA to use as features, other types of metabolic information on these tumors are lacking. However, using outputs from genome-scale metabolic models of these TCGA tumors would provide an additional source of metabolic features which could be used in the prediction of radiation response. We have previously shown that integration of gene expression and mutation information from TCGA patient tumors, as well as kinetic and thermodynamic parameters from publicly-available repositories, into personalized genome-scale FBA models of tumor metabolism can provide accurate predictions of differences in redox metabolism between radiation-sensitive and radiation-resistant tumors. This was accomplished by using objective functions which captured the reduction of redox cofactors and clearance of reactive oxygen species, which is necessary by radiation-resistant tumors to withstand oxidative stress caused by ionizing radiation therapy [196]. It is hypothesized that an objective function maximizing the production of a particular metabolite will provide accurate predictions of relative metabolite production between radiation-sensitive and radiation-resistant tumors if this metabolite is particularly beneficial to either class of tumors and thus the metabolic network of these tumors is optimized to produce elevated levels of this metabolite. These metabolites with large differential production between radiation-sensitive and -resistant tumors (and that would

be accurately predicted by FBA models) would be useful towards prediction of radiation sensitivity and thus selected by feature selection, whereas metabolites with similar production rates among the two classes (which may not be well captured by the objective function and thus not accurately predicted by FBA models) would not be useful and would be removed by feature selection.

To predict the production of individual metabolites within TCGA tumors for use in machine learning classifiers, a metabolome-wide objective function screen was performed, where the maximization of each metabolite was individually used as an objective function, and the resulting objective value was used as a prediction for maximum metabolite production (Figure 5-17). All 871 metabolites in Recon3D that (1) had KEGG database ID's, (2) were not present in the extracellular media, and (3) were capable of being produced by all FBA tumor models, were included in screen. Figure 5-17 shows the ratio of average predicted production of each metabolite in radiation-resistant versus radiation-sensitive tumors. There were a few major classes of metabolites which showed at least 2-times greater predicted production in radiation-resistant tumors. Sulfur-containing metabolites including lipoamide, a known antioxidant, and L-cystathionine, an intermediate in the synthesis of cysteine and glutathione, were among the metabolites with greatest differential production in radiation-resistant tumor models [317-320]. Additionally, the production of lipid and fatty acid metabolites including glycerol, γ-linolenic acid, and butyric acid was also predicted to be significantly higher in radiation-resistant tumors; these metabolites are involved in antioxidation, lipid peroxidation, and phospholipid production, all of which are implicated in the response to ionizing radiation [175, 179, 321-323]. These findings are consistent with the significant enrichment of genes

involved in sulfur metabolism, lipid metabolism, and oxidoreductase activity from the gene expression classifier for radiation response. Interestingly, the inflammation-mediating eicosanoids PGJ2, PGD2, and TXA2 showed increased predicted production in radiation-resistant tumors. Finally, while fewer metabolites were predicted to have at least 2-times greater production in radiation-sensitive tumors, many nucleotide metabolites, including thymine and deoxyribose, were among this group.

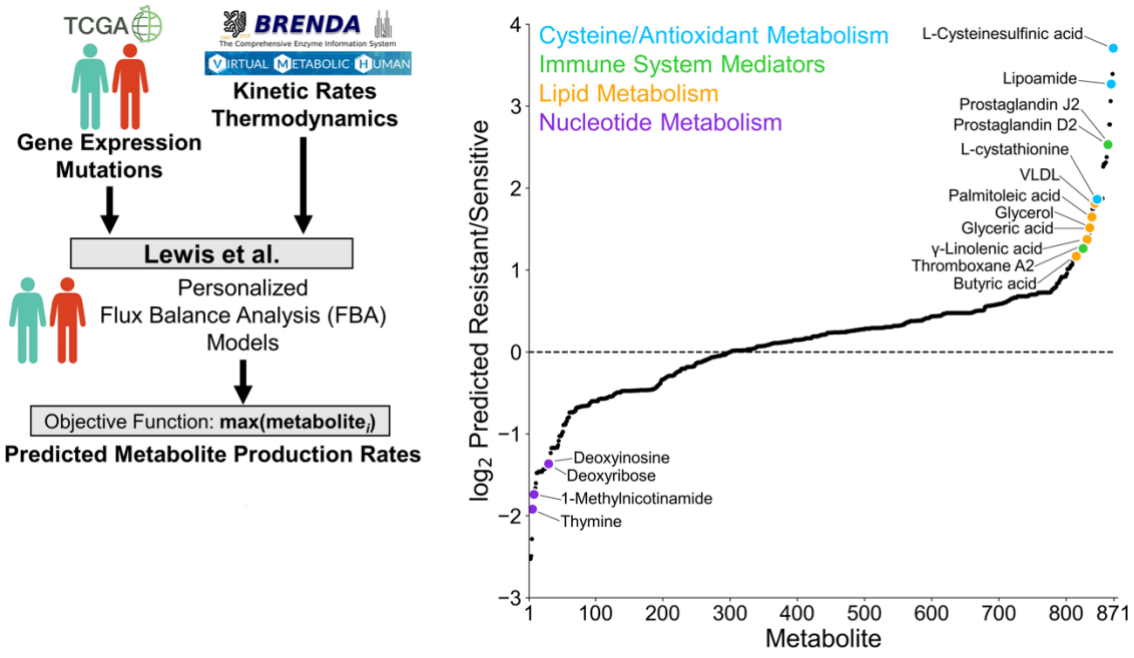


Figure 5-17. (Left) Schematic showing the integration of multi-omic data from TCGA and publicly-available repositories, as well as objective functions maximizing the production of individual metabolites, to predict relative metabolite production rates in TCGA patient tumors. (Right) Model-predicted production of 871 metabolites between radiation-sensitive and radiation-resistant tumors. Values are the $\log_2$ ratio of average production of the metabolite in radiation-resistant tumors divided by average production in radiation-sensitive tumors; $\log_2$(Predicted Resistant/Sensitive) < 0 signifies that a metabolite has greater maximum production

in radiation-sensitive tumors, $\log_2$(Predicted Resistant/Sensitive) > 0 signifies that a metabolite has greater maximum production in radiation-resistant tumors. Metabolites are rank ordered based on increasing ratio value. Colors indicate metabolite classes (cyan: sulfur-containing metabolites; green: eicosanoids; orange: lipids and fatty acids; purple: nucleotides).

To determine whether radiation-resistant cancers do show increased or decreased levels of these FBA-predicted classes of metabolites, an analysis of metabolomics data from the NCI-60 cancer cell line panel was performed (Figure 5-18) [324]. Among 139 metabolites screened, cholesterol was found to have the most positive correlation between metabolite concentration and cell line surviving fraction at 2 Gy radiation (SF2) within the NCI-60 panel. Cholesterol, closely related to many lipid metabolites predicted to have increased production in radiation-resistant tumors, was previously shown to be upregulated in response to ionizing radiation and may protect against lipid peroxidation [325, 326]. Glycerol and glyceric acid, other metabolites involved in lipid metabolism which were discovered from the FBA objective function screen, as well as glutathione, an important sulfur-containing antioxidant, also showed largely positive correlations with radiation response. Many uracil-containing metabolites as well as oritidine 5'-phosphate, a precursor to uridine monophosphate, were among the metabolites with largely negative correlations with radiation response, in agreement with FBA predictions of nucleotide production.

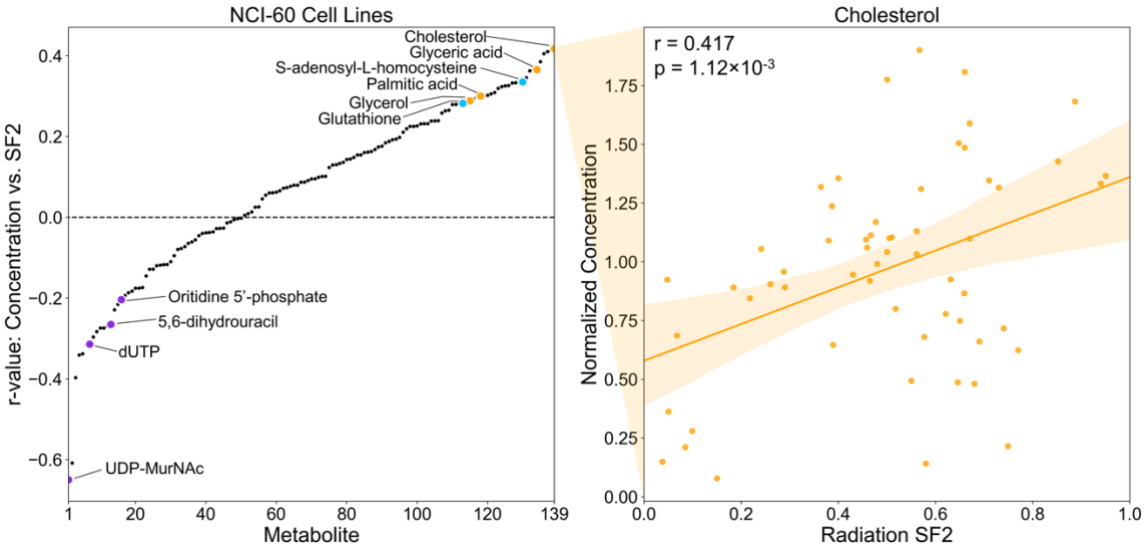


Figure 5-18. (Left) Correlation between metabolite concentration and surviving fraction at 2 Gy radiation (SF2) among 139 experimentally-measured metabolites in the NCI-60 panel of cancer cell lines. Metabolites are ranked by the correlation coefficient between concentration and SF2. Metabolite classes are colored as in (B). (Right) Example regression between concentration and SF2 for cholesterol.

Although the NCI-60 screen provided evidence of increased/decreased production of general classes of metabolites over a wide range of radiation-resistant cancers, not enough metabolites were available to compare model-predicted and experimentally-measured levels of individual metabolites. To this end, untargeted metabolomics was performed on four matched pairs of radiation-sensitive and radiation-resistant cell lines from four different cancer types (Figure 5-19) [175, 220, 327, 328]. For nucleotide, lipid, immune system-mediating, and cysteine/antioxidant metabolites that were identified in both the experimental metabolomics screen and the FBA-based objective function screen, the fold change in measured concentrations between matched radiation-resistant and -sensitive cell lines, as well as the average fold change across all four matched pairs, was compared to

the ratio of average predicted production in radiation-resistant versus radiation-sensitive FBA tumor models (Figure 5-19, Figure 5-20, Figure 5-21, Figure 5-22, Figure 5-23). FBA models correctly predicted that most nucleotide metabolites, including derivatives of adenine, guanine, thymine, and inosine, were downregulated in radiation-resistant cancers; additionally, cytidine derivatives were correctly predicted to be upregulated in radiation-resistant cancers, unlike most other nucleotides. Interestingly, there were a large number of lipid metabolites with greater experimentally-measured levels among the radiation-sensitive cell lines, as well as a large amount of heterogeneity between cell lines of different cancer types. Only two lipid metabolites (CDP-choline and succinate) had incorrect predictions in terms of matching signs of $\log_2$ ratio values with average experimentally-measured levels, and not having at least one cell line pair with a statistically-significant difference in the same direction as the FBA-predicted metabolite differences; this suggests that FBA models are able to capture the observed diversity in lipid levels. Although model-predictions of absolute oxidized (GSSG) and reduced (GSH) glutathione production did not match well with experimentally-measured values, previous model predictions of increased reduction potential of GSSG to GSH in radiation-resistant tumors agrees with experimental findings of greater GSH/GSSG ratios in radiation-resistant cell lines (Figure 4-13). Finally, model-predicted production of the antioxidant lipoamide, as well as immune mediators anandamide and 2-arachidonylglycerol, matched very well with experimental measurements, which were upregulated in nearly all radiation-resistant cell lines. Overall, these findings demonstrate that FBA-based genome-scale metabolic models provide accurate predictions of relative metabolite production between radiation-sensitive and -

resistant cancers, allowing for their use in machine learning classifiers for radiation sensitivity.

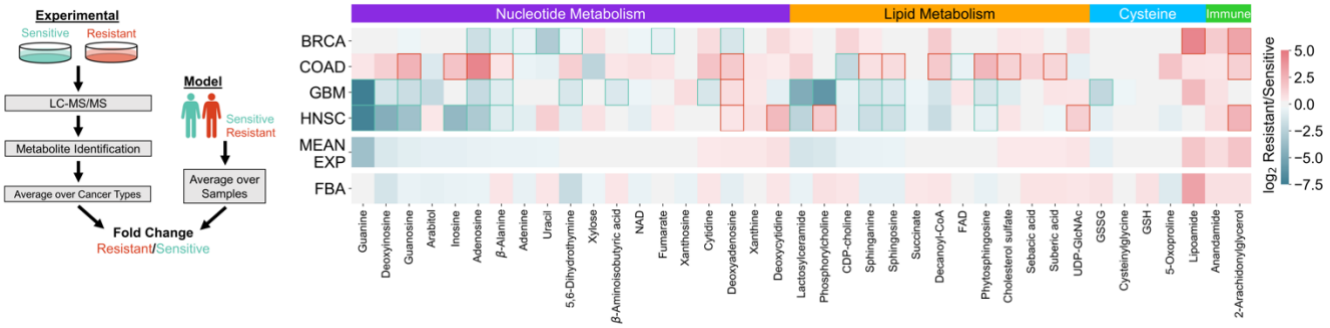


Figure 5-19. (Left) Schematic showing the comparison of model-predicted metabolite production in radiation-sensitive and -resistant TCGA tumors, with experimentally-measured metabolite concentrations in matched radiation-sensitive and -resistant cell lines from four different cancer types. (Right) Comparison of model-predicted and experimentally-measured ("FBA") metabolite levels, expressed as the $\log_2$ of the ratio of metabolite levels in radiation-resistant versus radiation-sensitive tumor models or cell lines. Experimentally-measured values are shown for each individual cell line pair, as well as averaged across all 4 cell line pairs ("MEAN EXP"). Statistically-significant differences in experimentally-measured metabolite levels within each cell line pair are represented by green (Sensitive > Resistant) or red (Resistant > Sensitive) box outlines. Metabolites are grouped by metabolite classes and ordered by experimentally-measured mean ratio values within each class.

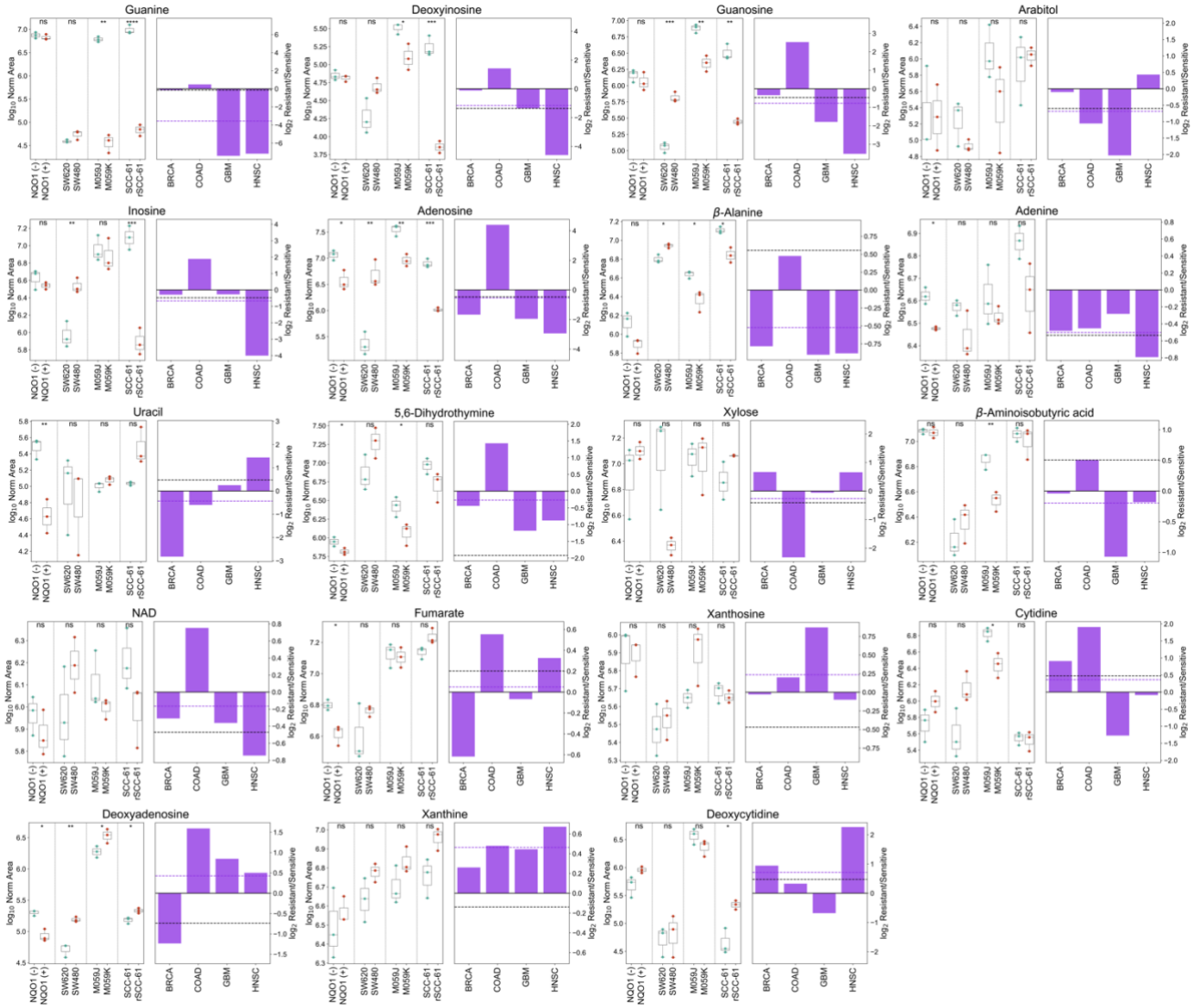


Figure 5-20. Experimentally-measured levels of nucleotide metabolites in radiation-sensitive and -resistant cell lines across four cancer types from untargeted metabolomics. (Left) Replicate metabolite concentrations from all four cell line pairs, with values expressed as the $\log_{10}$ of the normalized area from LC-MS/MS.

Statistical significance of the difference between radiation-sensitive and -resistant cell line values is shown.

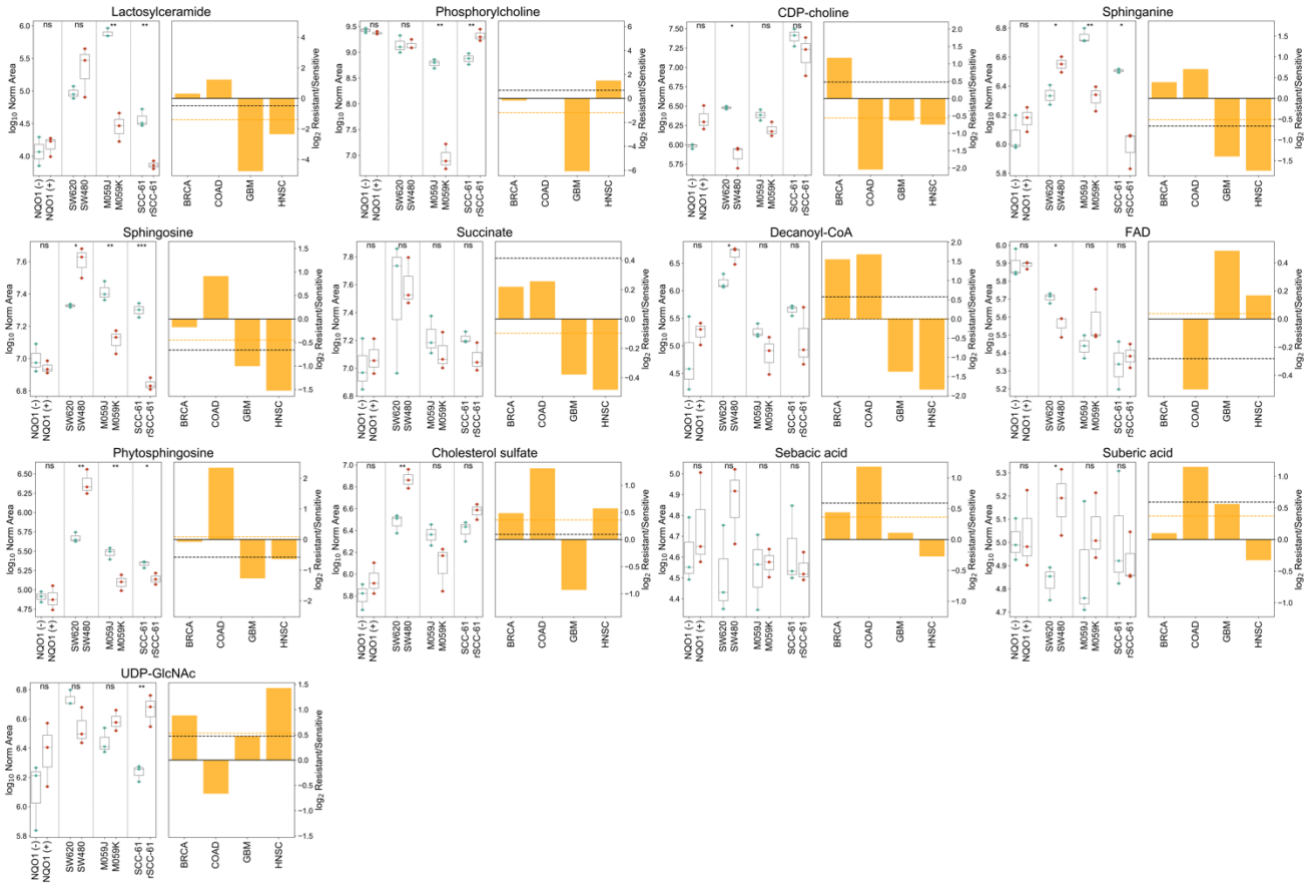


Figure 5-21. Experimentally-measured levels of lipid metabolites in radiation-sensitive and -resistant cell lines across four cancer types from untargeted metabolomics. (Left) Replicate metabolite concentrations from all four cell line pairs, with values expressed as the $\log_{10}$ of the normalized area from LC-MS/MS. Statistical significance of the difference between radiation-sensitive and -resistant cell line values is shown.

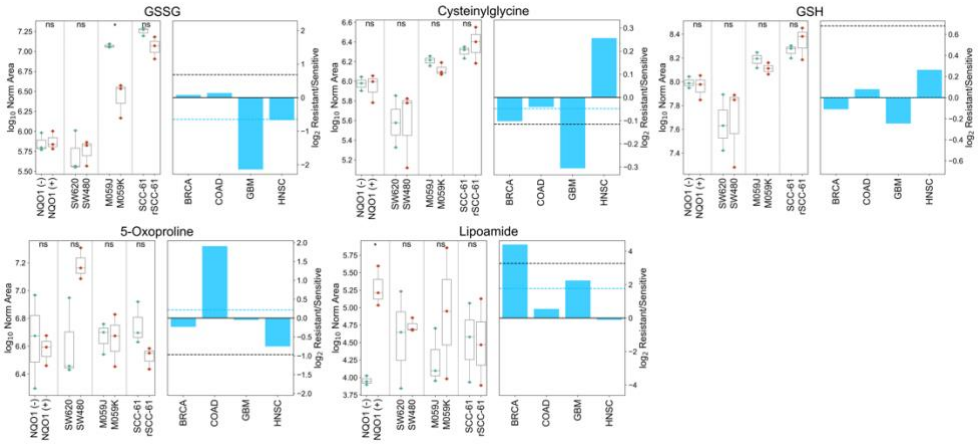


Figure 5-22. Experimentally-measured levels of cysteine/antioxidant metabolites in radiation-sensitive and -resistant cell lines across four cancer types from untargeted metabolomics. (Left) Replicate metabolite concentrations from all four cell line pairs, with values expressed as the $\log_{10}$ of the normalized area from LC-MS/MS. Statistical significance of the difference between radiation-sensitive and -resistant cell line values is shown.

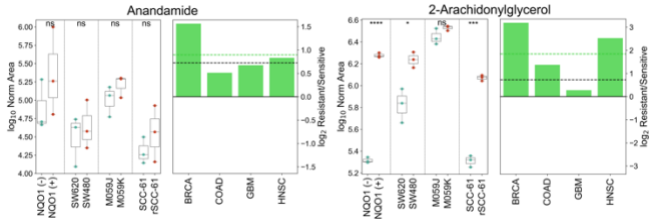


Figure 5-23. Experimentally-measured levels of immune system-mediating metabolites in radiation-sensitive and -resistant cell lines across four cancer types from untargeted metabolomics. (Left) Replicate metabolite concentrations from all four cell line pairs, with values expressed as the $\log_{10}$ of the normalized area from

LC-MS/MS. Statistical significance of the difference between radiation-sensitive and -resistant cell line values is shown.

5.3.4 Multi-omic classifier identifies clinical subpopulations of cancer patients

Using the dataset-independent ensemble architecture, a machine learning classifier integrating clinical, gene expression, mutation, and FBA-predicted metabolite production rates was developed. With an AUROC of 0.906 ± 0.004, this classifier has significantly greater performance compared to previously-developed machine learning classifiers for radiation response [46, 271]. Figure 5-24 shows that using a predicted probability p=0.5 threshold for classifying radiation-sensitive and -resistant tumors results in a much greater specificity than sensitivity, indicating that the classifier provides more accurate predictions for radiation-sensitive samples compared to radiation-resistant ones. Although this may be desired to give greater certainty for patients who are classified as radiation-resistant (small number of false positives), the classification threshold can be changed to provide a better balance between sensitivity and specificity.

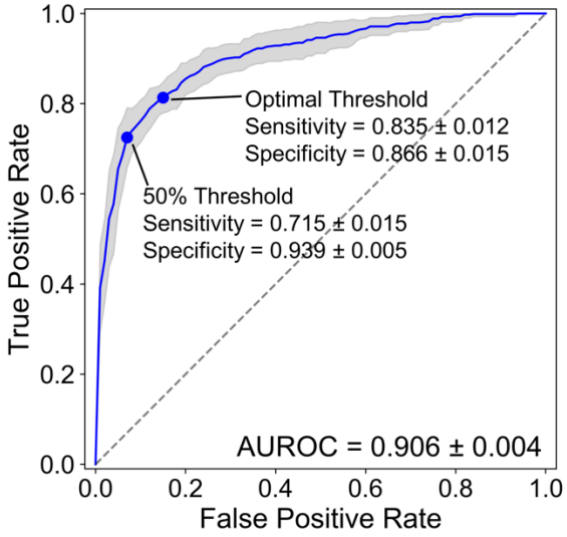


Figure 5-24. Receiver operating characteristic (ROC) curve, averaged across 20 training+validation/testing splits. Blue line represents the mean value across all splits, with the light blue error band representing +/- 1 standard deviation. The points on the ROC curve signifying both the p=0.5 threshold and optimal threshold (that which maximizes Youden's index) for differentiating radiation-sensitive and -resistant samples are shown, with associated sensitivity and specificity values.

725 of the 52,223 features from the four datasets (1.39%) were among the cumulative 95% sum of the classifier's mean absolute SHAP values (Figure 5-25). While the majority of these 725 features were gene expression (48.3%) and metabolite (32.6%) features, the clinical features contributed more than half of the total SHAP values (60.1%; Figure 5-26). Tumor histology had the largest overall impact on radiation response predictions; although there are known differences in radiation sensitivity between different cancer types, the differences in radiation response between histologic variants of the same cancer type is much less characterized [31, 329, 330]. Interestingly, the patient's response to different chemotherapies also had a significant impact; many of these chemotherapies

are radiation sensitizing and have known impacts on redox metabolism, including cisplatin, fluorouracil, cyclophosphamide, and carboplatin [121, 331-334]. Mutations with large feature importance scores included those directly involved in redox and lipid metabolism; IDH1 R132H (major SNP in the IDH1 gene) results in lower NADPH levels and the production of the oncometabolite 2-hydroxyglutarate, while BRAF V600E (major SNP in the BRAF gene) results in the upregulation of HMG-CoA lyase and subsequent conversion of HMG-CoA to the ketone body acetoacetate [215, 335].

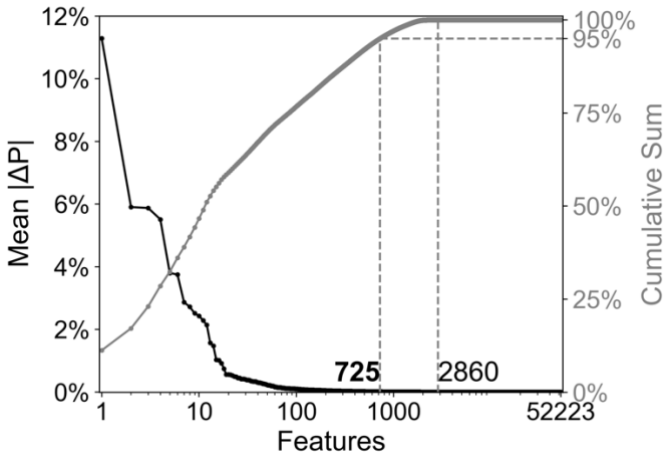


Figure 5-25. (Left; black overlay) Feature importance scores for individual features, calculated as the average absolute SHAP value (mean |ΔP|) across all tumor samples. Features are ordered by decreasing feature importance score. (Right; gray overlay) Cumulative sum of feature importance scores. Those features within the 95% cumulative sum of feature importance scores (bolded number) are kept for further analysis.

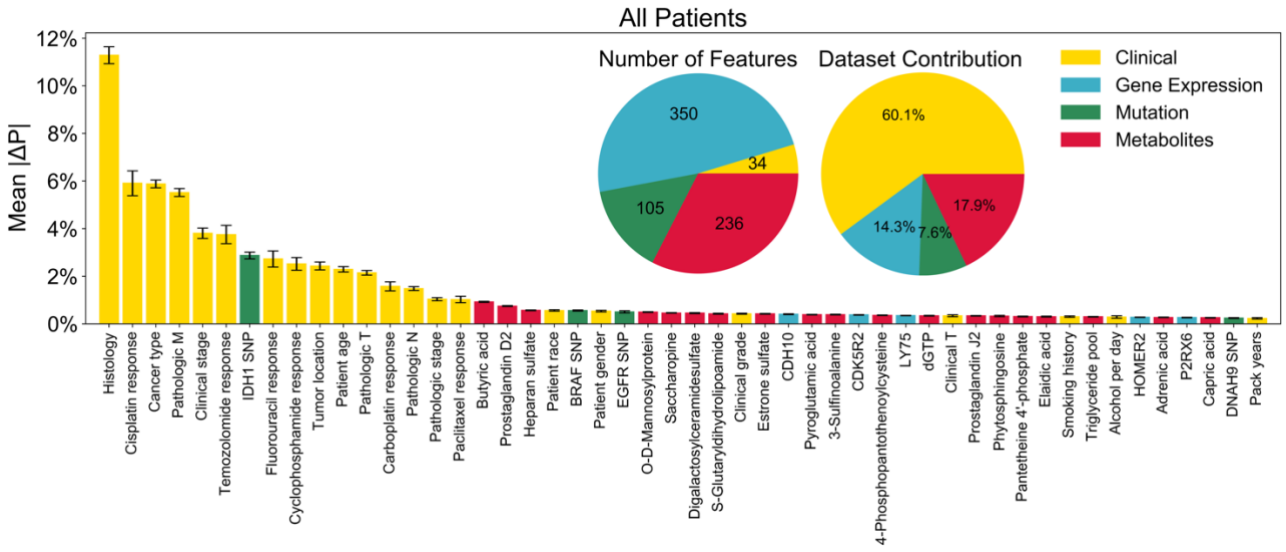


Figure 5-26. List of top 50 features with largest feature importance scores, colored based on their original dataset. (Inset, Left) Number of features contributing to the 725 total important features from each individual dataset. (Inset, Right) Relative contribution of features from each dataset to the sum of absolute SHAP values, averaged across all samples.

Averaged across all patients, clinical features supplied the majority of cumulative SHAP values and were most useful for prediction of radiation sensitivity. However, individual samples varied significantly in the contribution of clinical, gene expression, mutation, and metabolite features towards sensitivity classification (Figure 5-27). Using unsupervised clustering, three clusters of patients with varying contributions of clinical features were found (Figure 5-27, Figure 5-28). “High Clinical” patients are categorized by very large clinical feature contributions, meaning that performing multi-omic analysis of tumor biopsies from these patients to obtain gene expression, mutation, and metabolic features would not significantly improve the ability to classify these patients. On the other hand, “Low Clinical” patients have much lower clinical feature contributions; multi-omic

features provide the majority of cumulative SHAP values, with metabolic features alone providing nearly as much utility as clinical features for these patients (Figure 5-29). For this group of patients, while clinical features such as chemotherapeutic response have negligible utility, multi-omic measurements from tumor biopsies would be necessary to accurately predict response to radiation therapy. To determine what factors differentiate these three patient clusters, chi-squared tests comparing the expected and observed frequencies of patients with different values of each clinical factor were performed (Figure 5-30). Significant heterogeneity in clinical clusters was seen between patients of different cancer types; while some cancers including prostate (PRAD) and bladder (BLCA) cancer were categorized mainly by "Low Clinical" patients, others including breast (BRCA) and cervical (CESC) cancer consisted mostly of "High Clinical" patients (Figure 5-31). Histological subtypes of the same cancer type also showed significant heterogeneity, again demonstrating the importance of tumor histological characterization on radiation sensitivity prediction.

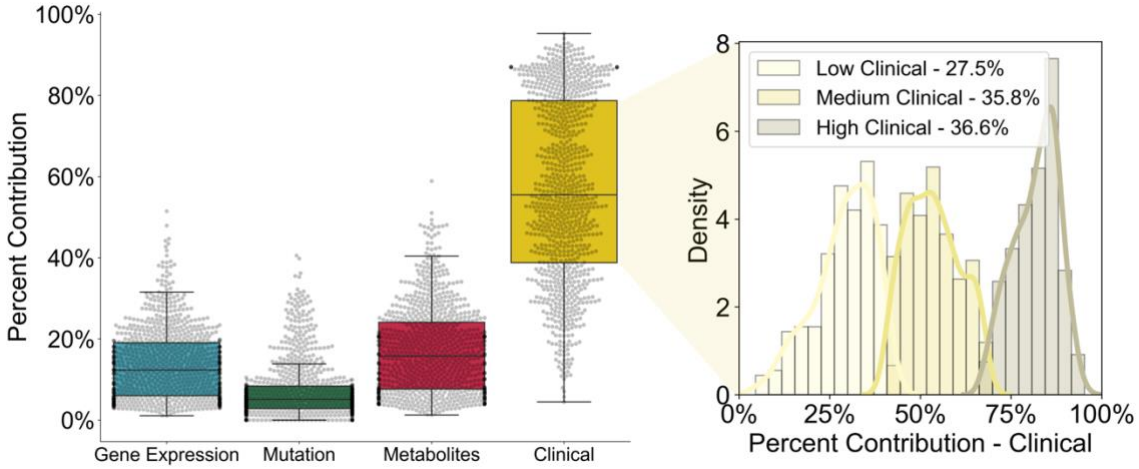


Figure 5-27. (Left) Relative contribution of features from each dataset to the sum of absolute SHAP values for each individual sample. (Right) Clustering of samples into

"Low", "Medium", and "High" clinical groups based on the relative contribution of features from the clinical dataset toward the sum of absolute SHAP values. The optimal number of clusters was calculated based on maximizing the gap statistic from k-means clustering (Figure 5-28).

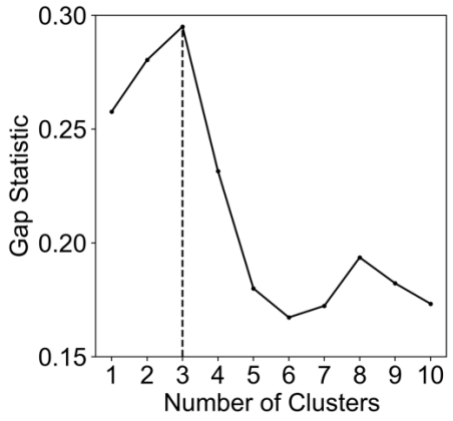


Figure 5-28. Gap statistic associated with k-means clustering of clinical dataset contributions for each sample from the multi-omic classifier, as a function of number of clusters (k). The optimal number of clusters is that which yields the largest gap statistic.

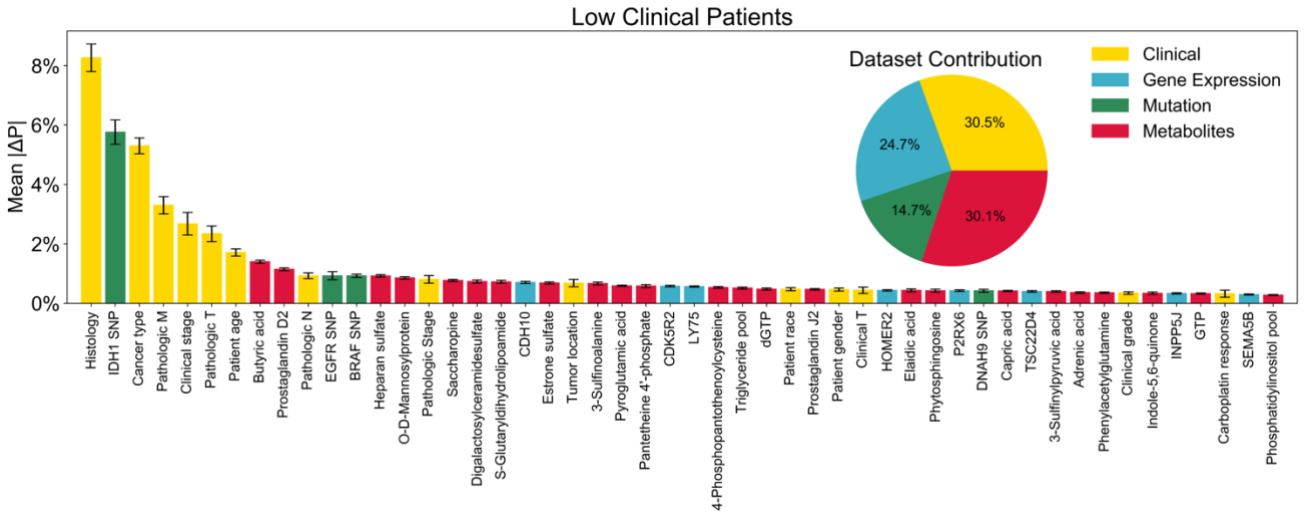


Figure 5-29. List of top 50 features with largest feature importance scores among samples within the "Low Clinical" cluster. (Inset) Relative contribution of features from each dataset to the sum of absolute SHAP values, averaged across all samples within the "Low Clinical" cluster.

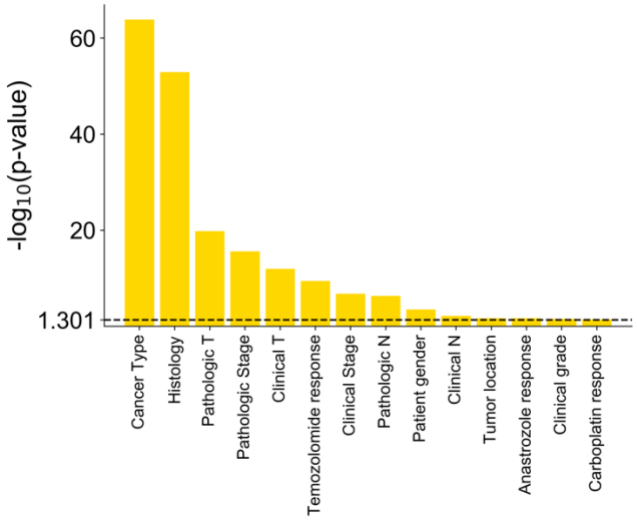


Figure 5-30. Statistical significance of patient clustering into "Low/Medium/High" clinical groups based on different clinical factors, as calculated by the chi-squared

test with Yates' correction. Only statistically significant ($p \leq 0.05$) clinical factors are shown.

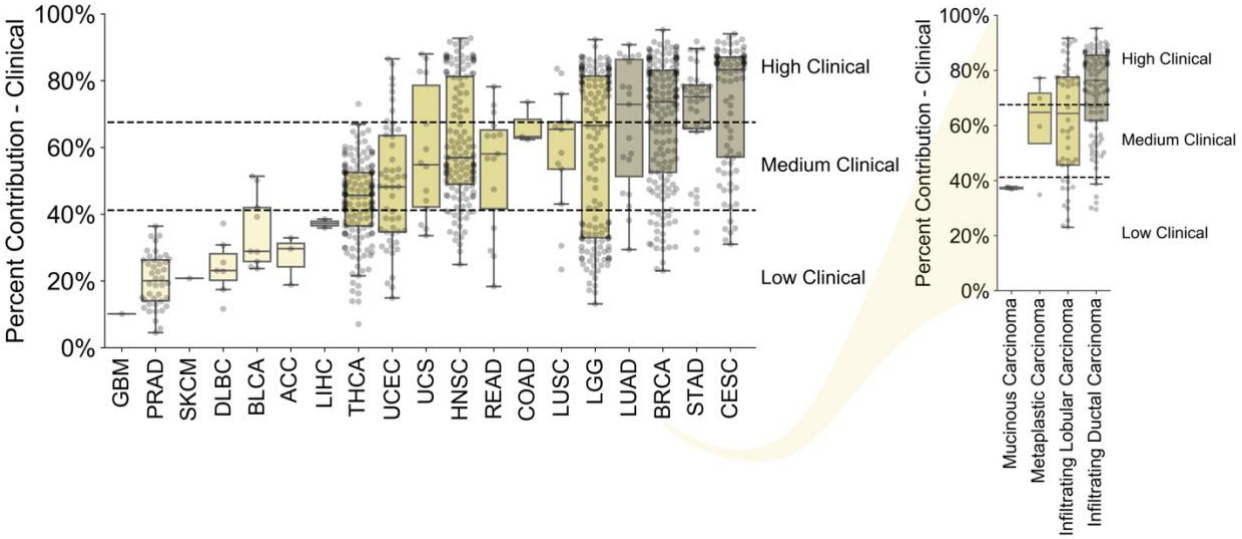


Figure 5-31. (Left) Heterogeneity of samples from different cancer types in terms of clinical dataset contribution and distribution into "Low/Medium/High" clinical groups. (Right) Heterogeneity of samples from different histological subtypes of breast cancer (BRCA) in terms of clinical dataset contribution and distribution into "Low/Medium/High" clinical groups.

The meta learner of the ensemble architecture outputs the probability that each dataset provides the most utility in classification, used as weights for the predicted probability of radiation resistance from each dataset-independent base learner. These weights provide accurate predictions of cumulative SHAP values of the clinical dataset relative to other datasets, and can differentiate between patients within the "Low Clinical" cluster and "Medium/High Clinical" clusters (Figure 5-32). Because the meta learner can be run independently of base learners on testing samples, it can be used to predict *a priori*

whether patient clinical information is sufficient to accurately predict radiation sensitivity, or whether multi-omic features from tumor biopsy samples are needed.

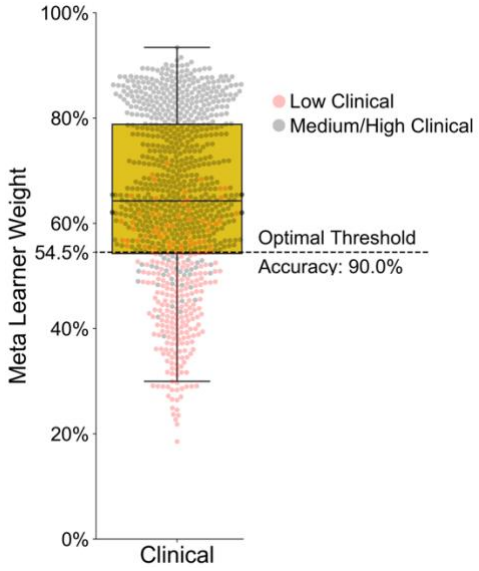


Figure 5-32. Using the meta learner weight for the clinical dataset (probability that the clinical dataset is the optimal dataset for each individual sample) to determine *a priori* clustering into the "Low Clinical" or "Medium/High Clinical" groups. Dotted line signifies the optimal threshold separating "Low Clinical" or "Medium/High Clinical" groups, which maximizes the accuracy of predicting clinical cluster based on meta learner weight.

5.3.5 Metabolic features highlight network-level involvement

Analysis of gene expression and multi-omic classifiers for radiation sensitivity suggests that cellular metabolism is integral to the tumor response to ionizing radiation, and that metabolic features can contribute significantly to the prediction of radiation response in individual patient tumors. The 236 FBA-based metabolic features with significant feature importance scores from the multi-omic classifier were examined to

identify individual metabolic biomarkers for radiation sensitivity. Metabolite set enrichment analysis indicates significant enrichment of several metabolic pathways involved in central carbon metabolism, lipid metabolism, and nucleotide metabolism (Figure 5-33). Interestingly, many metabolites are involved in transport between various cellular compartments as well as the extracellular space, suggesting that these biomarker metabolites could be measured in interstitial fluid or blood as well as within the tumor itself. To identify individual metabolites with the largest impact on radiation response, the spearman correlation between predicted metabolite production rate and SHAP value (change in probability of radiation resistance due to that metabolite) across all patients was calculated for each metabolite (Figure 5-34). These correlation values signify whether increased tumor metabolite levels are predicted to increase (positive correlation) or decrease (negative correlation) the probability of radiation resistance, and to what extent.

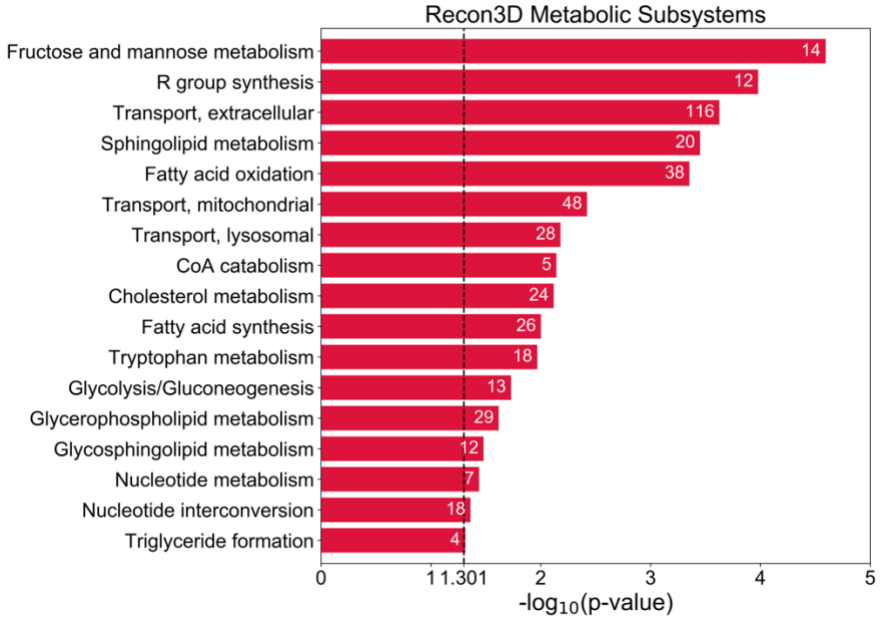


Figure 5-33. Metabolite-set enrichment analysis of the 236 significant metabolic features from the multi-omic classifier into Recon3D metabolic subsystems. The

statistical significance and number of metabolites in each subsystem are shown. Only statistically significant ($p \leq 0.05$) subsystems are shown.

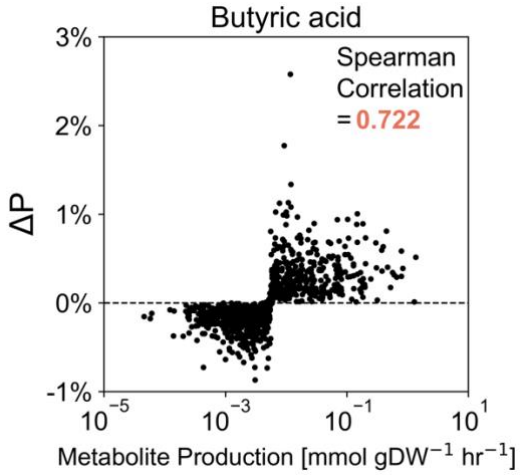


Figure 5-34. Regression between metabolic feature SHAP value and predicted metabolite production for a representative metabolite (butyric acid), showing values for each individual patient tumor, as well as the Spearman correlation coefficient.

Figure 5-35 highlights many of the significant metabolic features, their predicted impact on radiation response, and their interconnections with central carbon metabolism. Significant metabolites in the glycolysis and TCA pathways (fructose 1,6-bisphosphate, 3-phosphoglyceric acid, succinyl-CoA, and succinate) were all positively correlated with radiation resistance, while genes promoting gluconeogenesis (PCK2 and LDHA) were associated with radiation sensitivity. Fructose 2,6-bisphosphate, an allosteric regulator of PFK-1 that activates glucose breakdown, had one of the most positive correlation values with radiation resistance. Interestingly, many metabolites in early mannose metabolism had positive correlations values. Experimental studies have shown that radiation induces the upregulation of mannose-6-phosphate receptors as well as production of high mannose-

type N-glycans; however, the function of mannose in radiation response has yet to be determined [336, 337].

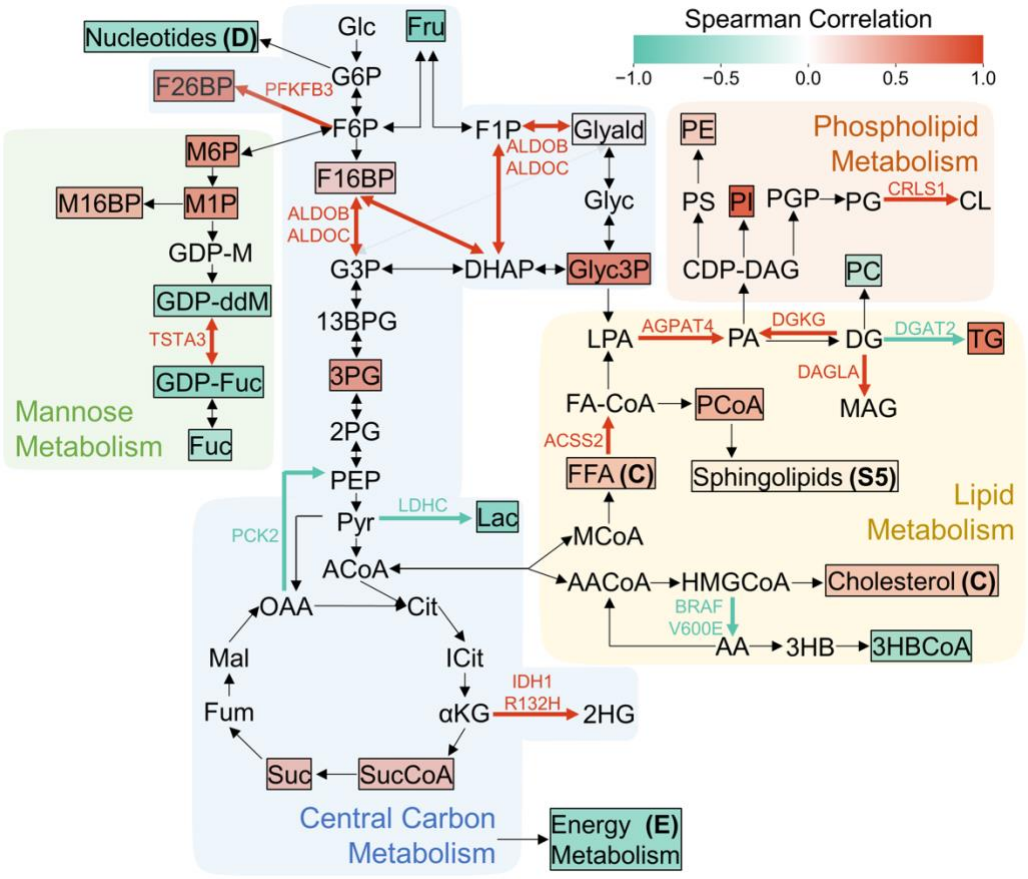


Figure 5-35. Overview of significant metabolic biomarkers and their interconnections within the human metabolic network. Different metabolic pathways are shown with colored backgrounds. Significant metabolic features are denoted by colored boxes, where the color indicates the Spearman correlation coefficient between the SHAP value for that feature and predicted metabolite production rate (see Figure 5-34); metabolites with positive correlations (red) are predicted to have higher production in radiation-resistant tumors, whereas metabolites with negative correlations (green) are predicted to have greater production in radiation-sensitive tumors. Significant gene expression and mutation

features are denoted by colored reaction arrows, either in red (higher gene expression or mutation rate in radiation-resistant tumors) or in green (higher gene expression or mutation rate in radiation-sensitive tumors).

Greater fluxes through central carbon metabolism in radiation-resistant tumor models resulted in increased production of both glycerol 3-phosphate and free fatty acids. This allowed for enhanced production of lipid and phospholipid metabolites, including phosphatidylinositol, phosphatidylethanolamine, and triglycerides. The majority of significant metabolites involved in fatty acid and cholesterol metabolism also had positive correlations with radiation resistance (Figure 5-36). Many of these metabolites have previously-identified roles as antioxidants with effects on lipid peroxidation, including capric acid, butyric acid, eicosatrienoic acid, docosapentaenoic acid, γ-linolenic acid, and myristic acid [338-343]. There was a dichotomy between sphingolipid metabolites that had either strongly positive or strongly negative correlations with radiation resistance; this agrees with previous studies showing that while the production of ceramides in response to DNA damage promotes apoptosis, other pro-survival sphingolipids can promote therapeutic resistance (Figure 5-37) [344, 345].

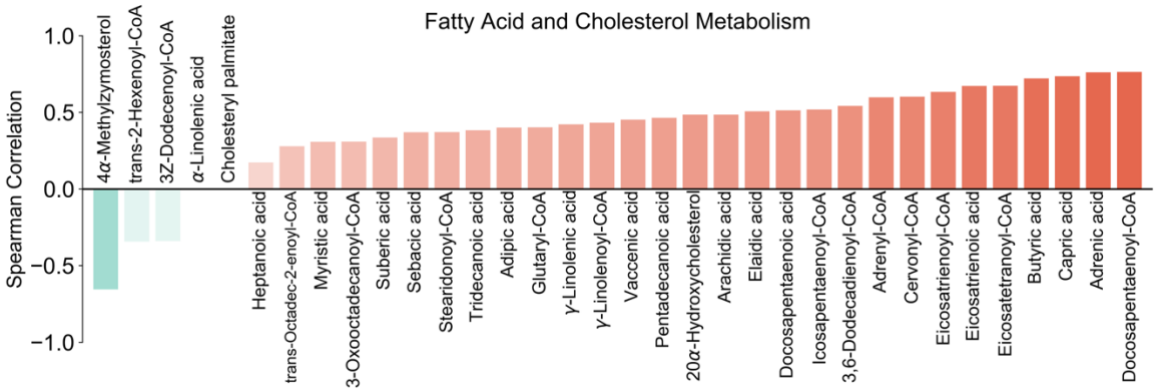


Figure 5-36. Spearman correlation coefficients of significant metabolic biomarkers involved in fatty acid and cholesterol metabolism.

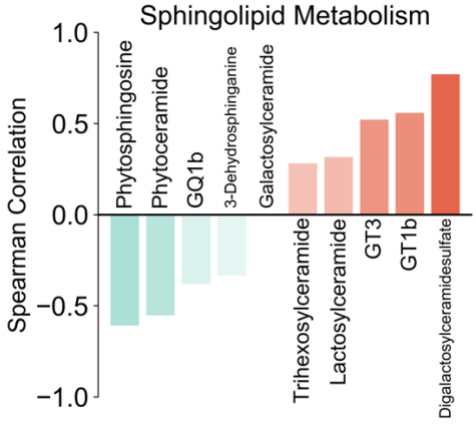


Figure 5-37. Spearman correlation coefficients of significant metabolic biomarkers involved in sphingolipid metabolism.

While metabolic fluxes through lipid and fatty acid-generating pathways are enhanced in radiation-resistant tumor models, model-predicted production of metabolites involved in nucleotide and energy metabolism, also closely connected to central carbon metabolism, was highly correlated with radiation sensitivity (Figure 5-38). Despite the expectation that radiation-resistant tumors would upregulate nucleotide synthesis for enhanced DNA damage repair, many of these nucleotide metabolites showed

downregulation experimentally in radiation-resistant cell lines (Figure 5-19). Previous studies found enzyme upregulation and increased production of pyrimidine nucleotides in radiation-resistant cancers, which was observed experimentally with cytidine and deoxycytidine in Figure 5-19 [346]. However, machine learning classifiers did not identify pyrimidines as having significant impact on predicting radiation response; instead, purine metabolites, which showed decreased production in radiation-resistant cancers from both FBA models and experimental metabolomics, did have significant impacts on machine learning predictions. The disparate metabolic pathways for generating these two nucleotide classes may allow for differences in regulation and function in response to ionizing radiation therapy [347, 348]. The observed negative correlation between ATP production and radiation resistance is likely due to purine synthesis rather than availability of high-energy phosphates; FBA models predict significantly greater capacity for conversion of ADP to ATP in radiation-resistant tumors, in accordance with previous experimental findings (Figure 5-39) [349, 350].

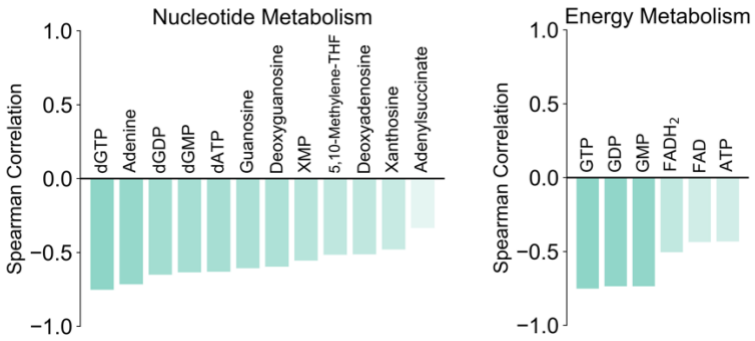


Figure 5-38. Spearman correlation coefficients of significant metabolic biomarkers involved in (left) nucleotide metabolism, and (right) energy metabolism.

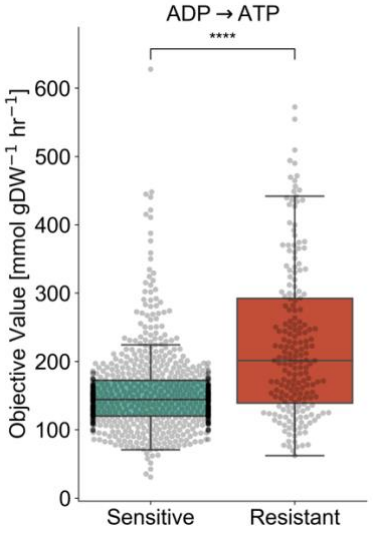


Figure 5-39. Maximal conversion of ADP to ATP in radiation-sensitive and -resistant tumors, predicted from FBA metabolic models.

Many eicosanoid precursors and inflammatory mediators were found to significantly impact classification and had positive correlations with radiation resistance (Figure 5-40). Increased eicosanoid production in radiation-resistant models was due to both elevated levels of arachidonic acid precursors (membrane phospholipids including dihomo-γ-linolenic acid) and increased expression of eicosanoid-producing enzymes. The majority

of significant sulfur-containing metabolites (66.7%) also had positive correlations with radiation resistance (Figure 5-41). However, among the 8 sulfur-containing metabolites in the top 50 of all features among “Low Clinical” samples, only half had positive correlation values. This suggests that the sulfur-containing functional groups such as thiol groups may play other roles in the biological response to ionizing radiation other than providing antioxidant function and therapeutic resistance [351]. Together, these findings indicate that metabolic features from multiple interconnected pathways including central carbon metabolism, lipid metabolism, and nucleotide metabolism are viable biomarkers for the prediction of radiation response.

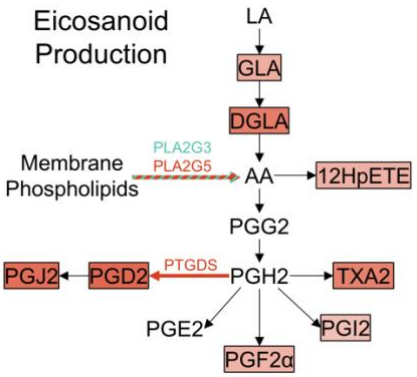


Figure 5-40. Metabolic pathway of eicosanoid production, highlighting significant metabolite features (colored boxes) and gene expression features (colored arrows).

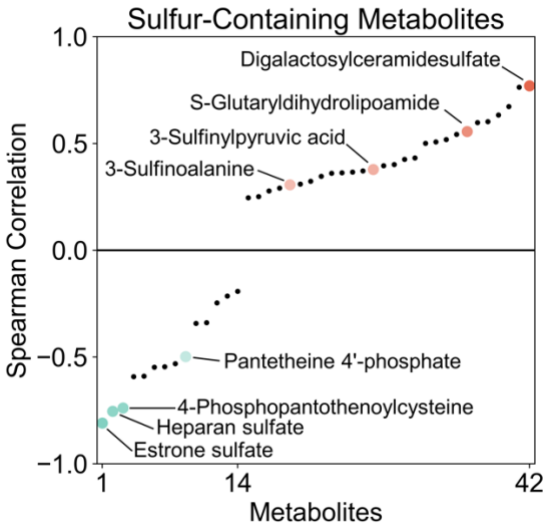


Figure 5-41. Spearman correlation coefficients of significant sulfur-containing metabolic biomarkers. Colored dots signify metabolites within the top 50 of all features among "Low Clinical" samples in the multi-omic classifier.

5.3.6 Personalized predictions of non-invasive clinical and metabolic biomarkers

While the integration of genome-scale mutational, transcriptomic, and metabolomic information into multi-omic classifiers may provide optimal performance in predicting radiation therapy response, it requires invasive and costly tests to measure these biomarkers from tumor biopsies. To determine the efficacy of a non-invasive predictor of radiation sensitivity, a machine learning classifier was developed which integrates clinical data derived from non-invasive means (excluding any staging or histological information from tumor biopsies) with FBA-predicted production rates of metabolites known to be quantifiable in human blood samples (Figure 5-42) [178]. Surprisingly, this non-invasive classifier showed similar performance to the multi-omic classifier, with increased sensitivity and decreased specificity (Figure 5-43). The higher sensitivity of the non-invasive classifier minimizes the number of false negatives and ensures that few radiation-

resistant patients will be missed, making it ideal as a first-line screening test [352]. If a patient tests positive, further analysis can be performed to obtain features needed for the multi-omic classifier, which has a higher specificity and is ideal as a second-line diagnostic test.

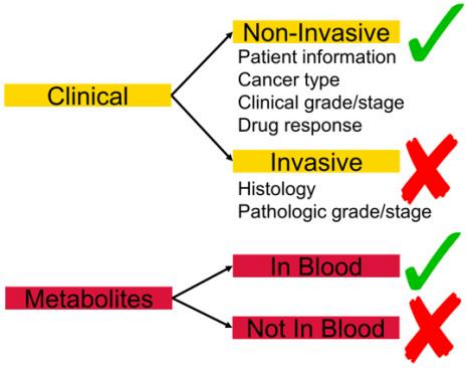


Figure 5-42. Inclusion of non-invasive clinical features and blood-based metabolic features, with exclusion of invasive clinical features and non-blood-based metabolic features.

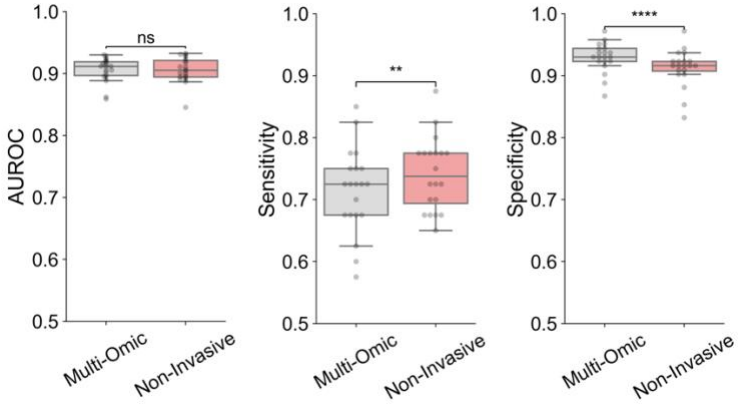


Figure 5-43. Comparison of model AUROC, sensitivity, and specificity between multi-omic and non-invasive classifiers.

97 of the 363 features from the two datasets (26.7%) were among the cumulative 95% sum of the classifier's mean absolute SHAP values (Figure 5-44). Similar to the multi-omic classifier, the individual patient contribution of clinical features towards cumulative SHAP values formed a multi-modal distribution; unsupervised clustering separated these patients into "Low Clinical" and "High Clinical" groups (Figure 5-45). While cancer type constitutes about one-fourth of the predictive utility among "Low Clinical" patients, all other non-invasive clinical factors cumulatively make up only about one-fourth as well (Figure 5-46). For these patients, blood metabolite features provide almost one-half of the cumulative SHAP values, demonstrating the value of blood-based biomarkers in radiation sensitivity prediction. Many of the major metabolite groups from the multi-omic classifier also represent the metabolites with largest mean SHAP values in the non-invasive classifier, including lipids (e.g. butyric acid), nucleotides (e.g. GDP), and immune system mediators (e.g. prostaglandins J2 and D2).

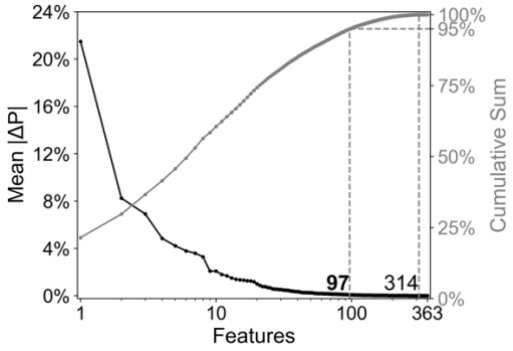


Figure 5-44. (Left; black overlay) Feature importance scores for individual features, calculated as the average absolute SHAP value (mean |ΔP|) across all tumor samples. Features are ordered by decreasing feature importance score. (Right; gray overlay) Cumulative sum of feature importance scores. Those features within the

95% cumulative sum of feature importance scores (bolded number) are kept for further analysis.

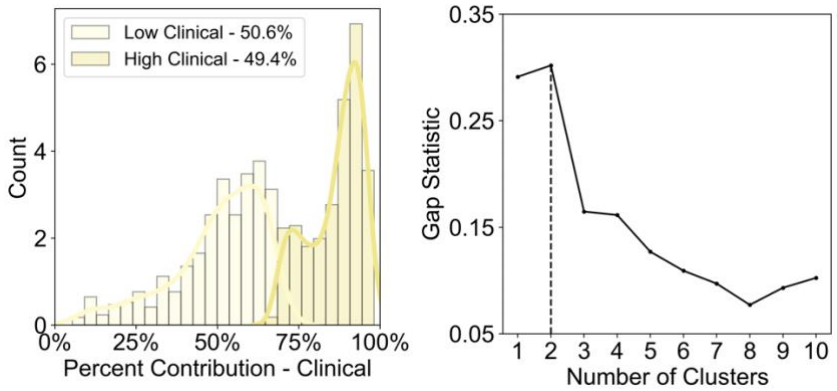


Figure 5-45. (Left) Clustering of samples into "Low" and "High" clinical groups based on the relative contribution of features from the clinical dataset toward the sum of absolute SHAP values. The optimal number of clusters was calculated based on maximizing the gap statistic from k-means clustering (Right).

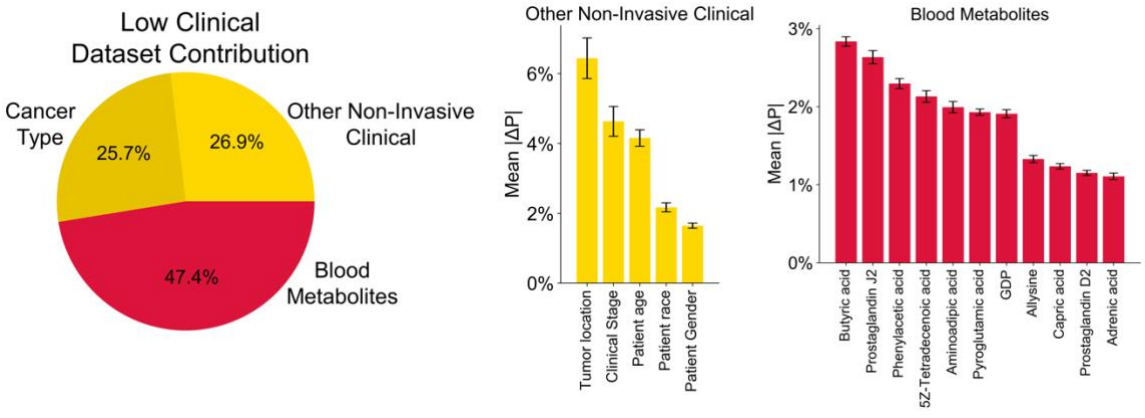


Figure 5-46. Analysis of clinical and metabolic feature contributions among the "Low" clinical group. Features with mean absolute SHAP values above 1% are explicitly shown.

Figure 5-47 provides a representative example of how clinical and metabolic features are cumulatively used by the non-invasive classifier to predict the posterior probability of radiation resistance in individual patients. For this patient, blood metabolic features constituted 64.5% of the cumulative SHAP values and increased the predicted probability of radiation resistance from 48.9% to 73.5%, allowing for the correct classification into the radiation-resistant class. This patient had the highest SHAP value among all patients for the metabolite 5Z-tetradecenoic acid, suggesting that this particular metabolite is a very useful blood-based biomarker for their diagnosis of radiation sensitivity. On the other hand, this patient had one of the lowest SHAP values for the metabolite phenylacetylglutamine, suggesting a lower relative utility of this biomarker compared to other cancer patients. Overall, this methodology allows for the determination of personalized biomarker panels with maximal diagnostic utility for individual patients, reducing the number of features needed to be measured per patient as well as decreasing associated costs for diagnosis.

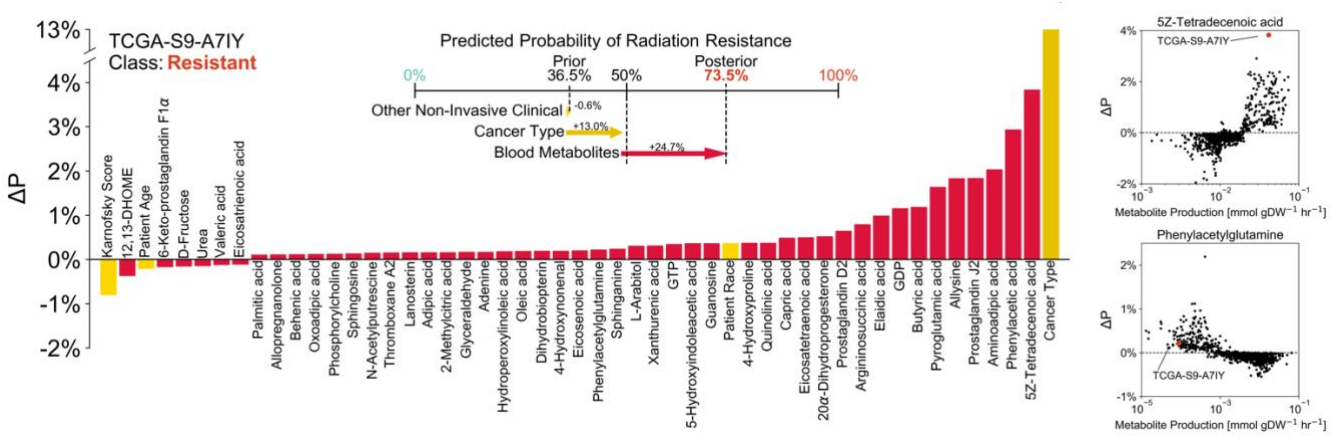


Figure 5-47. (Left) Breakdown of feature contributions towards prediction of radiation resistance in a representative radiation-resistant patient (TCGA-S9-A7IY). (Upper) Contribution of each data type towards the progression from prior to posterior probability of radiation resistance. (Lower) SHAP values of individual

features for this patient. (Right) Plots of feature SHAP value versus predicted metabolite production rate for two features, illustrating (upper) a feature whose absolute SHAP value for patient TCGA-S9-A7IY is larger than average, and (lower) a feature whose absolute SHAP value for patient TCGA-S9-A7IY is smaller than average.

5.4 Discussion

While interest in the prediction of radiation sensitivity in cancer patients has been present for decades, these predictors have yet to be used in the clinical setting [272]. The recent progression of high-throughput technologies including whole genome sequencing and RNA-seq have facilitated the generation of multi-omic datasets from cancer cell line and tumor samples, including the Cancer Cell Line Encyclopedia (CCLE) and The Cancer Genome Atlas (TCGA); these datasets provide tens of thousands of biological features for exploring the pathophysiological mechanisms of radiation response and the generation of predictive models [2, 307]. Development of sophisticated machine learning algorithms has simultaneously provided methods for analyzing these high-dimensional biological datasets and making reliable predictions from them [271]. Nevertheless, recently-developed classifiers for radiation response have failed to integrate multiple -omic modalities or utilize other modeling strategies for feature generation and biomarker identification. Here, we propose a novel strategy of integrating metabolic features from personalized genome-scale FBA models with clinical, genomic, and transcriptomic data to predict radiation sensitivity in individual TCGA patient tumors. By utilizing a dataset-independent ensemble architecture with gradient boosting models and Bayesian optimization, we obtain superior accuracy compared to previously-developed classifiers for radiation response.

Additionally, by integrating predicted metabolite production rates from FBA models, which were experimentally validated through untargeted metabolomics on matched radiation-sensitive and -resistant cancer cell lines, metabolic biomarkers were identified across the human metabolic network which correlate with radiation resistance. Finally, to our knowledge this is the first study to demonstrate the capability of identifying multi-omic and non-invasive biomarkers with maximum diagnostic utility for predicting radiation response in individual cancer patients.

Many of the identified metabolic biomarkers from this study agree with previous experimental evidence from radiation-resistant cancers and radiation-sensitizing therapies. Most of the identified fatty acid and cholesterol metabolite biomarkers had positive correlations with radiation resistance (Figure 5-36). Radiation-resistant head and neck cancer cells were shown to have enhanced fatty acid production from increased expression of fatty acid synthase [179]. Additionally, ionizing radiation was shown to cause increased cholesterol production in lung cancer cells [325]. Plasma levels of total and HDL cholesterol were found to be increased in radiation-resistant SPRET/EiJ mice compared to radiation-sensitive BALB/cByJ mice, implicating cholesterol as a potential non-invasive metabolic biomarker (Figure 5-48). Treatment with HMG-CoA reductase inhibitors including simvastatin have been shown to sensitize prostate cancer cells to radiation therapy, potentially by compromising DNA damage repair [353-357]. Other agreements between model predictions and experimental studies include the potential role of immune-modulating eicosanoids in radiation resistance. Several prostaglandins including PGD2, PGF2α, PGI2, and PGJ2, as well as thromboxane A2, had positive correlations with radiation resistance (Figure 5-40). Previous experimental studies have associated increased

production of prostaglandins PGF2α and PGE2 with radiation resistance, and have found that cyclooxygenase inhibitors such as aspirin may enhance response to radiation therapy and improve outcomes of cervical, prostate, and rectal cancer patients [358-363]. These findings suggest that lipid and eicosanoid metabolites may have utility as both diagnostic biomarkers and therapeutic targets for improving radiation sensitivity.

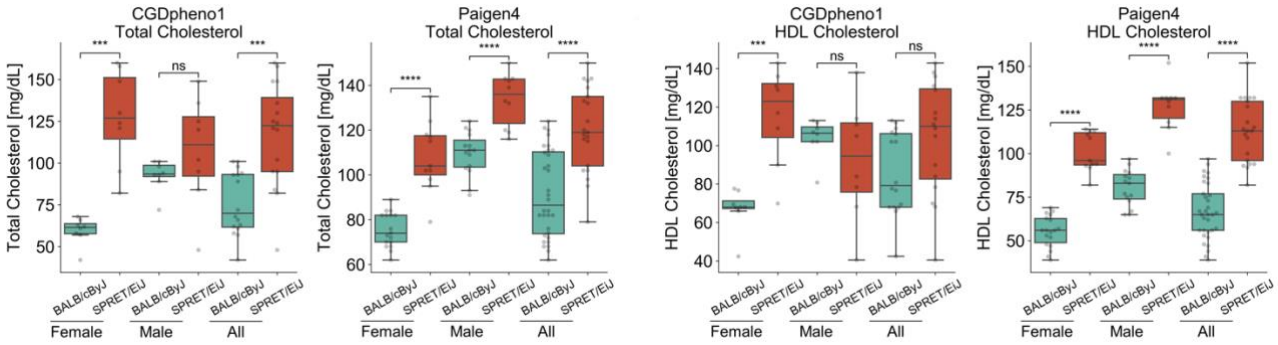


Figure 5-48. Comparison of (left) plasma total cholesterol levels, and (right) plasma HDL cholesterol levels, in radiation-sensitive BALB/cByJ mice versus radiation-resistant SPRET/EiJ mice from the CGDpheno1 and Paigen4 datasets in the Mouse Phenome Database.

Stacked generalization is a form of ensemble learning where the output of multiple machine learning classifiers ("base learners") is used as input for another classifier ("meta learner"), which determines the optimal combination of original predictions to use as the final class prediction [364]. This machine learning architecture has been shown to improve predictive accuracy compared to individual classifiers in multiple medical applications [365-369]. However, in most cases there is only one input dataset being supplied to the multiple base learners. Here, we instead input different -omics datasets to separate base learners, and use the meta learner to determine the relative contribution of the different

datasets to the overall class prediction. While it may be anticipated that this architecture would decrease performance accuracy since it prevents interactions between features from different datasets, we have shown that it improves performance across multiple metrics in the prediction of radiation response (Figure 5-6). One possible explanation is that this dataset-independent ensemble architecture can better accommodate multiple high-dimensional datasets (where the number of features is much larger than the number of samples) compared to concatenating these datasets into one with even higher dimensionality. We propose that this architecture may provide optimal performance in other situations where multiple high-dimensional datasets (such as multi-omics datasets) are being integrated, which is becoming increasingly common in biomedical applications [370]. Another benefit of this ensemble architecture is that the meta learner can be utilized independently of the base learners after training to predict which individual datasets will provide the most utility for the *a priori* diagnosis of radiation sensitivity in individual patients. For example, the meta learner can accurately differentiate between "Low Clinical" patients (with large contributions of gene expression, mutation, and metabolic datasets from patient biopsy samples and genome-scale metabolic modeling) and "High Clinical" patients (with greater contribution from clinical data from electronic medical records) (Figure 5-32). "Low Clinical" and "High Clinical" patients also differed significantly in the diagnostic utility of individual features, including chemotherapy response, IDH1 SNP, and lipid metabolite levels (Figure 5-29). Thus, the meta learner's ability to stratify patient populations allows for optimal planning and resource allocation for collecting biological measurements with diagnostic utility for individual cancer patients. Moreover, the use of gradient boosting models as the base and meta learners provides a significant amount of

embedded feature selection, demonstrated by the identification of only 725 (1.39%) out of 52,223 total features as important for class prediction in the multi-omic classifier (Figure 5-25). This decrease in model complexity not only lowers the cost of measuring biological features needed for prediction, but also improves the interpretability of models which improves the likelihood of adoption by clinicians [371].

In addition to demonstrating the utility of multi-omic data for the classification of radiation response, we showed that a classifier utilizing non-invasive clinical information and blood-based metabolic biomarkers can predict radiation sensitivity with comparable accuracy (Figure 5-43). Blood-based diagnostic tools are garnering attention for their use in early detection, monitoring, and optimal treatment identification for cancer patients [372]. An area where the identification of novel circulating biomarkers could provide significant utility is in adaptive radiotherapy, where a cancer patient's radiation treatment is modified over time in response to the observed efficacy of previous treatments [373, 374]. The integration of machine learning and genome-scale metabolic modeling could be used to identify readily accessible blood-based metabolic signatures for both measuring the biological response of radiation therapy as well as identifying necessary changes or additions of radiation-sensitizing chemotherapies in response to changes in radiation sensitivity [375]. Both machine learning and genome-scale metabolic modeling are powerful tools for understanding cancer pathophysiology, identification of clinical biomarkers, and prediction of patient outcomes. Their synergistic integration will inevitably yield additional insights for improving precision medicine and long-term care of cancer patients.

CHAPTER 6. FBA MODELS OF RADIATION-SENSITIZING CHEMOTHERAPEUTICS

6.1 Introduction

A large number of chemotherapeutics have been found to increase sensitivity of cancer cells to ionizing radiation therapy through one or more of their mechanisms of action [376]. Thus, these drugs can be used in combination to enhance tumor cell death, or to decrease side effects on normal tissues by using lower radiation doses. Because ionizing radiation causes cellular death mainly through direct and indirect ROS-mediated DNA damage, many chemotherapies which either induce additional DNA damage or limit DNA damage repair mechanisms in response to radiation treatment have been found to have radiosensitizing effects, such as 5-fluorouracil and gemcitabine [96, 333]. Many chemotherapies that act in part by redox-based mechanisms of action are also radiation sensitizers; however, the impact of these drugs on oxidative stress and redox metabolism have seldom been associated with their radiation-sensitizing activity [6]. For example, while doxorubicin is known to undergo one-electron redox cycling resulting in NADPH consumption and superoxide generation, its radiosensitizing activity is mainly attributed to its effects on DNA damage caused by topoisomerase II inhibition and DNA intercalation [123-125, 377, 378]. Additionally, although cisplatin can cause superoxide generation through disruption of oxidative phosphorylation, radiation sensitization has instead been previously attributed to inhibition of nonhomologous end joining [106, 379]. Mechanistic modeling of redox-directed chemotherapeutic metabolism may provide insights into

possible mechanisms of radiation sensitization caused by altered redox biology and metabolism.

Kinetic ordinary differential equation (ODE)-based representations have traditionally been used to model the metabolism of redox-directed drugs. Finn et al. developed kinetic models of doxorubicin metabolism in acute lymphoblastic leukemia (ALL) cell lines with varying drug sensitivities, and showed alterations in cytotoxicity from switching between superoxide generation and NADPH depletion [125]. Similar kinetic models have been developed for cisplatin and cyclophosphamide metabolism [380, 381]. While kinetic models can more accurately capture the dynamic nature of drug metabolism, genome-scale metabolic models may present an alternative strategy for capturing the interconnections between drug metabolism and the human metabolic network which arise from use of common metabolic cofactors such as ATP, NADH, and NADPH; however, few genome-scale metabolic models of xenobiotic metabolism have been developed. Sahoo et al. incorporated a drug module composed of 210 metabolites and 721 metabolic reactions into the Recon2 metabolic reconstruction and defined drug metabolic objective functions to simulate the interconnection of many commonly-used medications (including statins, acetaminophen, and cyclosporine) with genome-scale metabolism [382]. Other studies have used genome-scale metabolic models of cancer cell lines to predict selective drug targets by identifying genes essential for cellular biomass production [141, 383]. Despite the advantage of genome-scale models in incorporating the thousands of human metabolic reactions involved in production or consumption of redox cofactors, FBA models of redox-directed chemotherapeutics and their interconnections with the redox metabolic network have yet to be developed.

Because chemotherapeutics have multiple complex mechanisms of action, potential modelling of these drugs using FBA may be best accomplished by maximizing multiple different objective functions that reflect these disparate mechanisms. FBA objective functions are often used to reflect the biological "objective" of the organism or cell type under study; for example, biomass production (proliferation) is often maximized in models of microorganisms and cancer cells [138]. Nonetheless, the choice of the optimal objective function for a particular application is often subjective, as a precise biological "objective" is unlikely to be captured by a single function. Few studies have investigated the use of multiple objective functions for obtaining better correspondence between model predictions and experimental validation. Garcia Sanchez et al. used a combination of compartment-specific objective functions to reflect differential compartmentalization of metabolic optimization observed in *Saccharomyces cerevisiae* [384]. Budinich et al. developed multi-objective flux balance analysis (MO-FBA) to reflect differences in objective functions between single microbial strains in an ecosystem with nutrient exchange between strains [385]. Despite these examples, there is currently a lack of systematic approaches for identifying optimal multi-feature objective functions for FBA models.

To investigate the role that genome-scale redox metabolism has on redox-driven chemotherapeutic activity, the mechanism of action of three commonly used chemotherapeutics (cisplatin, cyclophosphamide, and doxorubicin) were incorporated into FBA models of TCGA tumors. The metabolism of each of these three drugs involves one or multiple redox cofactors, including NADH, NADPH, and GSH (Figure 2-5, Figure 2-6, Figure 2-7). An optimal multi-feature objective function for each drug module was

identified by developing a novel objective function screening algorithm, comparing model-predicted objective values between drug-sensitive and drug-resistant TCGA tumors. These optimal objective functions were then used to predict chemotherapeutic response and radiation-sensitizing effect in radiation-resistant TCGA tumors, in order to determine in which patients these chemotherapies could be effectively used either alone or in combination with radiation therapy.

6.2 Methods

6.2.1 Integration of chemotherapeutic metabolism modules into FBA models

6.2.1.1 Prediction of kinetic rates of transport reactions

Matsson et al. previously quantified the kinetic rates of transport reactions for a range a substrates through a number of solute carrier (SLC) family and ATP-binding cassette (ABC) family transporters [386]. For a particular transport reaction in any of the drug modules (detailed below), if an exact match between both transporter and substrate was made with the Matsson et al. dataset, then that recorded kinetic rate was used. If a match was made between transporters but not substrate, a weighted average of the recorded kinetic rates for that transporter was used, with weights equal to the Tanimoto coefficient between the desired model substrate and Matsson et al. dataset substrate. If a match was made between substrates but not transporters, an average of the recorded kinetic rates for that substrate for all SLC-family transporters (for import reactions) or ABC-family transporters (for export reactions) was used. If there was no available match for transporter or substrate, both averages for transporters and weighted averages for substrates were performed as detailed above.

6.2.1.2 Cisplatin module

The following metabolites involved in cisplatin metabolism were added:

1. Cisplatin ("cis[]"): extracellular, cytosolic, mitochondrial
2. Cisplatin-mitochondrial DNA ("cis-mtDNA[]"): mitochondrial
3. Cisplatin-GSH ("cisGSH[]"): extracellular, cytosolic
4. Cisplatin-metallothionein ("cisMT[]"): extracellular, cytosolic

The following reactions involved in cisplatin metabolism were added:

- Exchange reactions:
    1. Exchange of cisplatin:

       1 cis[e] ←→ Ø, GPR: None, EC: None
    2. Exchange of cisplatin-GSH:

       1 cisGSH[e] ←→ Ø, GPR: None, EC: None
    3. Exchange of cisplatin-metallothionein:

       1 cisMT[e] ←→ Ø, GPR: None, EC: None
- Transport reactions:
    1. Transport of cisplatin from extracellular to cytosol:

       1 cis[e] → 1 cis[c], GPR: 1317, EC: *Transport (see above)*
    2. Transport of cisplatin from cytosol to extracellular:

       1 cis[c] + 1 atp[c] + 1 h2o[c] → 1 cis[e] + 1 adp[c] + 1 pi[c], GPR: 9429 or 538 or 540, EC: *Transport*
    3. Transport of cisplatin-GSH from cytosol to extracellular:

       1 cisGSH[c] + 1 atp[c] + 1 h2o[c] → 1 cisGSH[e] + 1 adp[c] + 1 pi[c], GPR: 1244, EC: *Transport*

4. Transport of cisplatin-metallothionein from cytosol to extracellular:

   1 cisMT[c] ←→ 1 cisMT[e], GPR: None, EC: None

5. Transport of cisplatin from cytosol to mitochondrion:

   1 cis[c] → 1 cis[m], GPR: 10063, EC: 1.9.3.1

- Cisplatin mechanism

  1. Cisplatin to cisplatin-mtDNA:

     1 cis[m] + 0.5 o2[m] → 1 cis-mtDNA[m] + 0.5 o2s[m], GPR: 4535 or 4536 or 4537 or 4538 or 4539 or 4540 or 4541 or 4519, EC: 1.9.3.1

- Cisplatin clearance

  1. Cisplatin to cisplatin-GSH:

     1 cis[c] + 1 gthrd[c] → 1 cisGSH[c] + 1 h[c], GPR: 2944 or 9446 or 2950 or 2952, EC: 2.5.1.18

  2. Cisplatin to cisplatin-metallothionein:

     2 cis[c] + 2 gthrd[c] → 1 cisMT[c] + 1 gthox[c] + 2 h[c], GPR: 4489 or 4502 or 4504 or 84560, EC: 1.8.5.1

- Artificial sinks

  1. Cisplatin-mtDNA sink:

     1 cis-mtDNA[m] → Ø, GPR: None, EC: None

6.2.1.3 Cyclophosphamide module

The following metabolites involved in cyclophosphamide metabolism were added:

1. Cyclophosphamide (“cpa[]”): extracellular, cytosolic
2. Dechloroethylcyclophosphamide (“dce-cpa[]”): extracellular, cytosolic

3. Chloroacetaldehyde (“caa[]”): extracellular, cytosolic
4. 4-Hydroxycyclophosphamide (“4oh-cpa[]”): cytosolic
5. 4-Ketocyclophosphamide (“4k-cpa[]”): extracellular, cytosolic
6. Aldophosphamide (“aldph[]”): cytosolic
7. Alcophosphamide (“alcph[]”): extracellular, cytosolic
8. Carboxyphosphamide (“cph[]”): extracellular, cytosolic
9. Phosphoramide mustard (“phmust[]”): cytosolic
10. Acrolein (“acr[]”): cytosolic
11. Phosphoramide mustard-GSH (“phmust-GSH[]”): extracellular, cytosolic
12. Acrolein-GSH (“acr-GSH[]”): extracellular, cytosolic

The following reactions involved in cyclophosphamide metabolism were added:

- Exchange reactions:
    1. Exchange of cyclophosphamide:

       1 cpa[e] ←→ Ø, GPR: None, EC: None
    2. Exchange of dechloroethylcyclophosphamide:

       1 dce-cpa[e] ←→ Ø, GPR: None, EC: None
    3. Exchange of chloroacetaldehyde:

       1 caa[e] ←→ Ø, GPR: None, EC: None
    4. Exchange of 4-ketocyclophosphamide:

       1 4k-cpa[e] ←→ Ø, GPR: None, EC: None
    5. Exchange of alcophosphamide:

       1 alcph[e] ←→ Ø, GPR: None, EC: None
    6. Exchange of carboxyphosphamide:

1 cph[e] ←→ Ø, GPR: None, EC: None

7. Exchange of phosphoramide mustard-GSH:

   1 phmust-GSH[e] ←→ Ø, GPR: None, EC: None

8. Exchange of acrolein-GSH:

   1 acr-GSH[e] ←→ Ø, GPR: None, EC: None

- Transport reactions:

1. Transport of cyclophosphamide from extracellular to cytosol:

   1 cpa[e] → 1 cpa[c], GPR: None, EC: None

2. Transport of cyclophosphamide from cytosol to extracellular:

   1 cpa[c] + 1 atp[c] + 1 h2o[c] → 1 cpa[e] + 1 adp[c] + 1 pi[c], GPR: 5243 or 9429, EC: *Transport*

3. Transport of dechloroethylcyclophosphamide from cytosol to extracellular:

   1 dce-cpa[c] ←→ 1 dce-cpa[e], GPR: None, EC: None

4. Transport of chloroacetaldehyde from cytosol to extracellular:

   1 caa[c] ←→ 1 caa[e], GPR: None, EC: None

5. Transport of 4-ketocyclophosphamide from cytosol to extracellular:

   1 4k-cpa[c] ←→ 1 4k-cpa[e], GPR: None, EC: None

6. Transport of alcophosphamide from cytosol to extracellular:

   1 alcph[c] ←→ 1 alcph[e], GPR: None, EC: None

7. Transport of carboxyphosphamide from cytosol to extracellular:

   1 cph[c] ←→ 1 cph[e], GPR: None, EC: None

8. Transport of phosphoramide mustard-GSH from cytosol to extracellular:

   1 phmust-GSH[c] + 1 atp[c] + 1 h2o[c] → 1 phmust-GSH[e] + 1 adp[c] + 1 pi[c], GPR: 1244, EC: *Transport*

9. Transport of acrolein-GSH from cytosol to extracellular:

   1 acr-GSH[c] + 1 atp[c] + 1 h2o[c] → 1 acr-GSH[e] + 1 adp[c] + 1 pi[c], GPR: 1244, EC: *Transport*

- Cyclophosphamide mechanism
  1. Cyclophosphamide to 4-hydroxycyclophosphamide:

     1 cpa[c] + 1 o2[c] + 1 nadph[c] + 1 h[c] → 1 4oh-cpa[c] + 1 h2o[c] + 1 nadp[c], GPR: 1555 or 1559 or 1557 or 1576, EC: 1.14.14.1

  2. 4-hydroxycyclophosphamide to aldophosphamide:

     1 4oh-cpa[e] ←→ 1 aldph[c], GPR: None, EC: None

  3. Aldophosphamide to phosphoramide mustard:

     1 aldph[c] → 1 phmust[c] + 1 acr[c], GPR: None, EC: None

- Cyclophosphamide clearance
  1. Cyclophosphamide to dechloroethylcyclophosphamide:

     1 cpa[c] + 1 o2[c] + 1 nadph[c] + 1 h[c] → 1 dce-cpa[c] + 1 caa[c] + 1 h2o[c] + 1 nadp[c], GPR: 1576 or 1577, EC: 1.14.14.1

  2. 4-hydroxycyclophosphamide to 4-ketocyclophosphamide:

     1 4oh-cpa[c] + 1 nad[c] → 1 4k-cpa[c] + 1 nadh[c] + 1 h[c], GPR: 124 or 125 or 126 or 127 or 128 or 130 or 131, EC: 1.1.1.1

  3. Aldophosphamide to alcophosphamide:

1 aldph[c] + 1 nadph[c] + 1 h[c] → 1 alcph[c] + 1 nadp[c], GPR: 10327 or 231 or 57016, EC: 1.1.1.21

4. Aldophosphamide to carboxyphosphamide:

   1 aldph[c] + 1 nad[c] + 1 h2o[c] → 1 cph[c] + 1 nadh[c] + 1 h[c], GPR: 216 or 218 or 7915, EC: 1.2.1.3

5. Phosphoramide mustard to phosphoramide mustard-GSH:

   1 phmust[c] + 1 gthrd[c] → 1 phmust-GSH + 1 h[c], GPR: 2938, EC: 2.5.1.18

6. Acrolein to acrolein-GSH:

   1 acr[c] + 1 gthrd[c] → 1 acr-GSH + 1 h[c], GPR: 2938, EC: 2.5.1.18

6.2.1.4 Doxorubicin module

The following metabolites involved in doxorubicin metabolism were added:

1. Doxorubicin ("doxQ[]"): extracellular, cytosolic
2. Doxorubicin semiquinone ("doxSQ[]"): cytosolic
3. Doxorubicinol ("doxHQ[]"): extracellular, cytosolic
4. 7-Deoxydoxorubicinone ("doxdeoxy[]"): extracellular, cytosolic
5. Daunosamine ("daun[]"): extracellular, cytosolic

The following reactions involved in cyclophosphamide metabolism were added:

- Exchange reactions:

  1. Exchange of doxorubicin:

     1 doxQ[e] ←→ Ø, GPR: None, EC: None

  2. Exchange of doxorubicinol:

     1 doxHQ[e] ←→ Ø, GPR: None, EC: None

3. Exchange of 7-deoxydoxorubicinone:

   1 doxdeoxy[e] ←→ Ø, GPR: None, EC: None

4. Exchange of daunosamine:

   1 daun[e] ←→ Ø, GPR: None, EC: None

- Transport reactions:

1. Transport of doxorubicin from extracellular to cytosol:

   1 doxQ[e] → 1 doxQ[c], GPR: 85413, EC: *Transport*

2. Transport of doxorubicin from cytosol to extracellular, RALBP1:

   1 doxQ[c] → 1 doxQ[e], GPR: 10928, EC: *Transport*

3. Transport of doxorubicin from cytosol to extracellular, ABC transporters:

   1 doxQ[c] + 1 atp[c] + 1 h2o[c] → 1 doxQ[e] + 1 adp[c] + 1 pi[c], GPR: 5243 or 4363 or 1244 or 10257 or 9429, EC: *Transport*

4. Transport of doxorubicinol from cytosol to extracellular:

   1 doxHQ[c] ←→ 1 doxHQ[e], GPR: None, EC: None

5. Transport of 7-deoxydoxorubicinone from cytosol to extracellular:

   1 doxdeoxy[c] ←→ 1 doxdeoxy[e], GPR: None, EC: None

6. Transport of daunosamine from cytosol to extracellular:

   1 daun[c] ←→ 1 daun[e], GPR: None, EC: None

- Doxorubicin mechanism

1. Doxorubicin to doxorubicin semiquinone, NOS3:

1 doxQ[c] + 2 o2[c] + 1 arg_L[c] + 1.5 nadph[c] + 1.5 h[c] → 1 doxSQ[c] + 2 h2o[c] + 1 no[c] + 1 citr_L[c] + 1.5 nadp[c], GPR: 4846, EC: 1.14.13.39

2. Doxorubicin to doxorubicin semiquinone, POR:

   1 doxQ[c] + 0.5 nadph[c] + 0.5 h[c] → 1 doxSQ[c] + 0.5 nadp[c], GPR: 5447, EC: 1.6.2.4

3. Doxorubicin semiquinone to doxorubicin:

   1 doxSQ[c] + 1 o2[c] → 1 doxQ[c] + 1 o2s[c] + 1 h[c]: GPR: None, EC: None

- Doxorubicin clearance

  1. Doxorubicin to doxorubicinol:

     1 doxQ[c] + 1 nadph[c] + 1 h[c] → 1 doxHQ[c] + 1 nadp[c], GPR: 873 or 874 or 10327 or 8644, EC: 1.1.1.2 or 1.1.1.184

  2. Doxorubicin to 7-deoxydoxorubicinone:

     1 doxQ[c] + 0.5 nadph[c] + 1.5 h[c] → 1 doxHQ[c] + 1 daun[c] + 0.5 nadp[c], GPR: 5447, EC: 1.6.2.4

6.2.2 Multi-feature objective function screen

For each drug module, the following were considered as possible single objective functions: (1) maximizing the production of each metabolite involved in the drug mechanism (including cofactors such as ATP and NADPH); (2) maximizing the flux through each reaction involved in the drug mechanism; and (3) maximizing the conversion between cofactor forms involved in the drug mechanism (e.g. conversion of ADP to ATP). The performance of each objective function was assessed by calculating the area under the

receiver operating characteristic curve (AUROC) for classifying drug-sensitive and drug-resistant TCGA tumors using the calculated objective value for each sample as threshold values; objective functions with the largest AUROC value provide optimal separation between drug-sensitive and drug-resistant classes.

Figure 6-1 provides an overview of the objective function screen to find the optimal multi-feature objective function for each drug module. All 2-feature objective functions composed of individual features with AUROC > 0.5 were considered. 10 combinations of weights for each feature of the 2-tuple were assessed. Weight values for each feature were randomly sampled from a uniform distribution centered at the inverse of the average single-feature objective value across all samples and spanning ± 2 standard deviations; this results in all feature objective values being on roughly equal orders of magnitude, allowing for accurate comparison of multiple-feature objective functions. After assessing each 2-feature objective function, those with AUROC greater than the best single-feature objective AUROC at any weight combination were kept for the next round. All the individual features among the top 2-feature objective functions were added systematically to the top 2-feature objective functions to create the 3-feature objective functions that will be assessed. This iterative pipeline continues until a round where no n-feature objective functions had an AUROC value greater than the top (n-1)-feature objective function. Afterwards, the top 10 overall objective functions underwent 100 iterations of Bayesian optimization to find the optimal weight combination which maximizes AUROC. The objective function and weights with the largest overall AUROC were used for prediction of drug metabolism response.

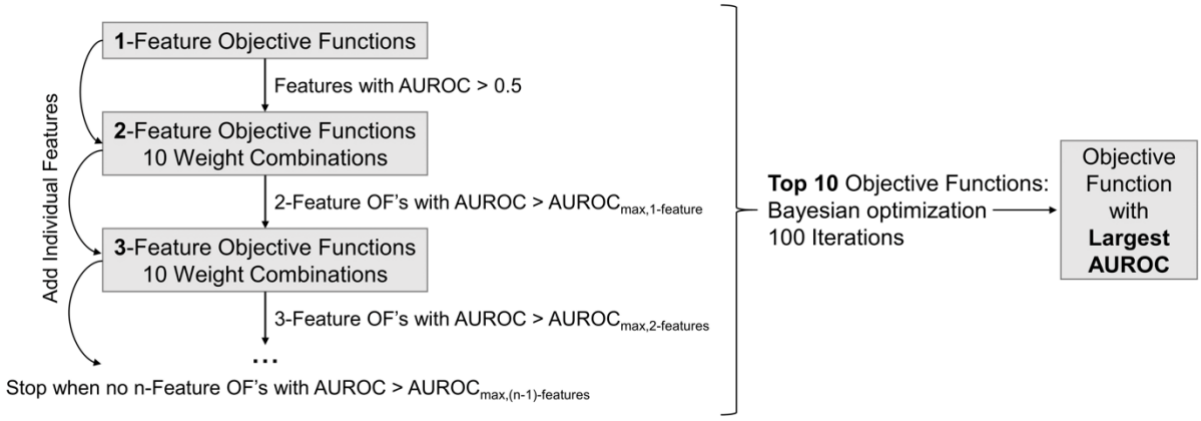


Figure 6-1. Overview of the objective function screen pipeline.

6.2.3 Machine learning regressors for radiation sensitizing effect

Implementation of machine learning regressors was analogous to that discussed in Chapter 5, except that base learners were performing regression tasks with mean squared error as the performance metric to be minimized:

$$MSE = \frac{1}{N_s}\sum_{i=1}^{N_s}\left(y_i^{true} - y_i^{pred}\right)^2 \tag{35}$$

where y_{itrue} is the true value for sample i, y_{ipred} is the predicted value for sample i, and N_s is the total number of samples.

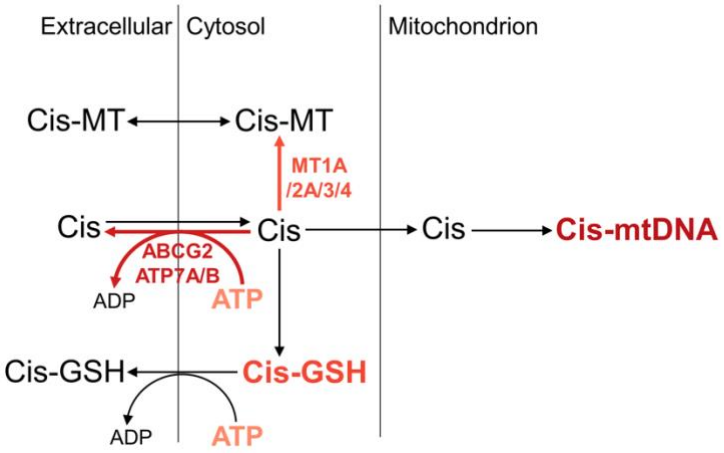
6.2.4 Code availability

Code for the simulation of chemotherapeutic response in FBA models of TCGA tumors is available at https://github.com/kemplab/FBA-chemosim. Jupyter notebooks are available for 1) implementation of multi-feature objective function screen; 2) simulation of chemotherapeutic response in FBA models; 3) assessment of model predictions.

6.3 Results

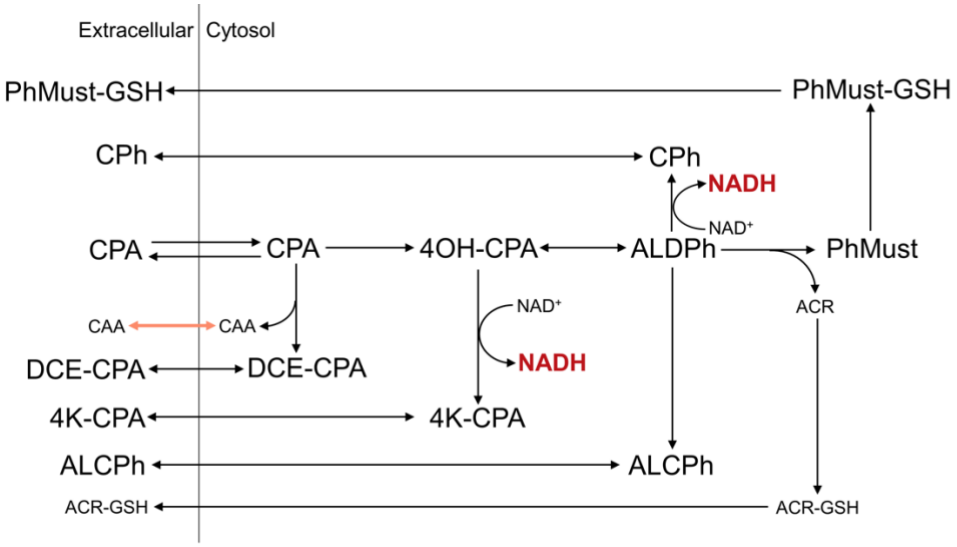
6.3.1 Optimal FBA objective functions utilize redox cofactors

Figure 6-2, Figure 6-3, and Figure 6-4 show the optimal objective functions for each drug module from the multi-feature objective function screen. All three optimal objective functions were composed of more than one feature. Two of the three optimal objective functions contained drug efflux reactions and/or the production of ATP necessary for drug efflux, consistent with previously-identified roles of drug efflux pumps including ABC transporters in chemotherapeutic resistance [387]. Additionally, two optimal objective functions contained drug clearance reactions or the production of metabolic cofactors including GSH and NADH which are utilized for drug clearance. Most notably, all three optimal objective functions contained metabolic reactions which utilize redox cofactors or the direct production of redox cofactors: cisplatin metabolism involves the production of the glutathionylated biproduct Cis-GSH, cyclophosphamide metabolism involves the direct production of NADH, and doxorubicin metabolism involves the NADPH-consuming reactions NOS3 and POR. The reduced forms of all three redox cofactors showed upregulated in radiation-resistant TCGA tumor models, suggesting that their utilization in chemotherapeutic metabolism may contribute to the drugs' radiation-sensitizing effect (Figure 4-12) [207, 388-390].



Rel. Weight	Feature
2.55	Cis-mtDNA Prod.
2.27	Cis Export
1.75	Cis-GSH Prod.
1.67	Cis → Cis-MT
1	ATP Production

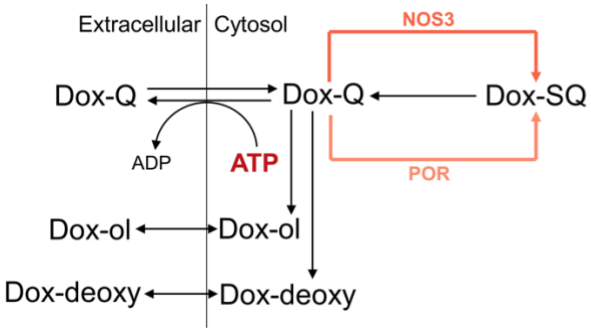
Figure 6-2. Optimal objective function for cisplatin response, highlighting the features involved (bolded metabolites and reaction arrows) and their relative weights. Features are colored based on relative weight values (largest value = dark red, lowest value = light orange).



Rel. Weight	Feature
1.40	NADH Production
1	CAA Export

Figure 6-3. Optimal objective function for cyclophosphamide response, highlighting the features involved (bolded metabolites and reaction arrows) and their relative

weights. Features are colored based on relative weight values (largest value = dark red, lowest value = light orange).



Rel. Weight	Feature
7.36	ATP Production
2.91	NOS3 Reaction
1	POR Reaction

Figure 6-4. Optimal objective function for doxorubicin response, highlighting the features involved (bolded metabolites and reaction arrows) and their relative weights. Features are colored based on relative weight values (largest value = dark red, lowest value = light orange).

Figure 6-5 shows the improvement in AUROC of each multi-feature objective function compared to the top single-feature objective function for each drug module; the DeLong test, a non-parametric approach for comparing the areas under two correlated ROC curves, shows that each of these three improvements is statistically significant [391]. For each drug module, the optimal objective value threshold for separating model-predicted drug-sensitive and -resistant classes was defined as that which maximizes Youden's J statistic:

$$J = \text{Sensitivity} + \text{Specificity} - 1 \tag{36}$$

This optimal threshold represents the point on the ROC curve which has the smallest distance from the top-left corner (where TPR=1 and FPR=0). Figure 6-6 shows the optimal objective value thresholds and corresponding Youden's J statistic for each drug module.

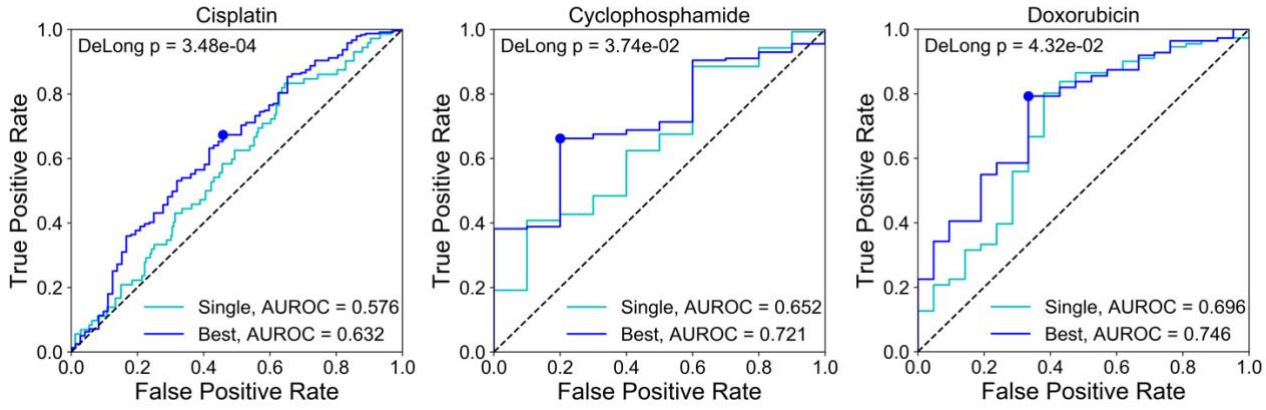


Figure 6-5. Comparison of ROC curves and AUROC values between the top single-feature objective function (light blue) and top multi-feature objective function (dark blue) for each drug module. The p-value of the DeLong test comparing the two AUROC values is shown. Dark blue dot represents the point on the ROC curve which maximizes Youden's J statistic.

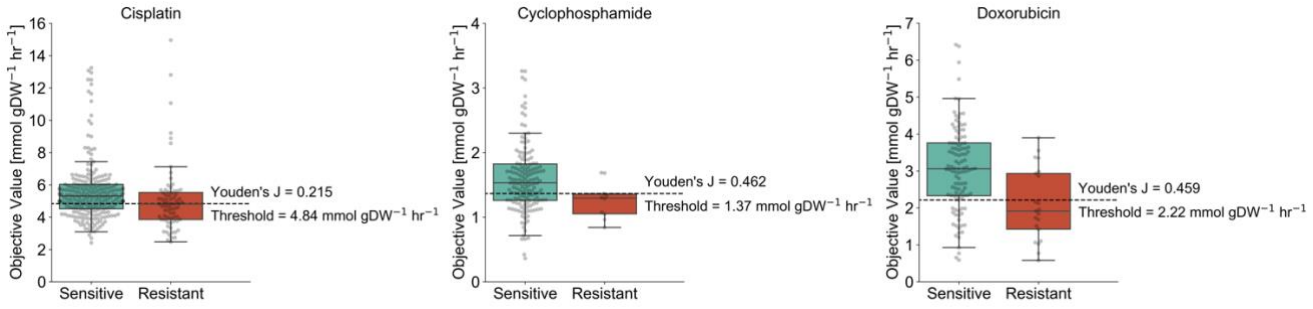


Figure 6-6. Separation of model-predicted drug-sensitive and drug-resistant classes based on the optimal objective value threshold which maximizes Youden's J statistic.

6.3.2 FBA models predict chemotherapy response and radiation sensitization

Optimal multi-feature objective functions and objective value thresholds for each drug module were used to predict chemotherapeutic response in radiation-resistant tumor models and determine for which radiation-resistant patients these alternative therapies may be beneficial (Figure 6-7). While 53.8% of radiation-resistant patients were predicted to be cisplatin-sensitive, only 40.7% were predicted to be cyclophosphamide-sensitive and 38.7% doxorubicin-sensitive. This larger proportion of radiation-resistant drug-resistant patients may be attributed to the similar metabolic effects of these chemotherapies as ionizing radiation therapy, in that they deplete cellular stores of reduced redox cofactors and promote ROS-mediated DNA damage [377, 392].

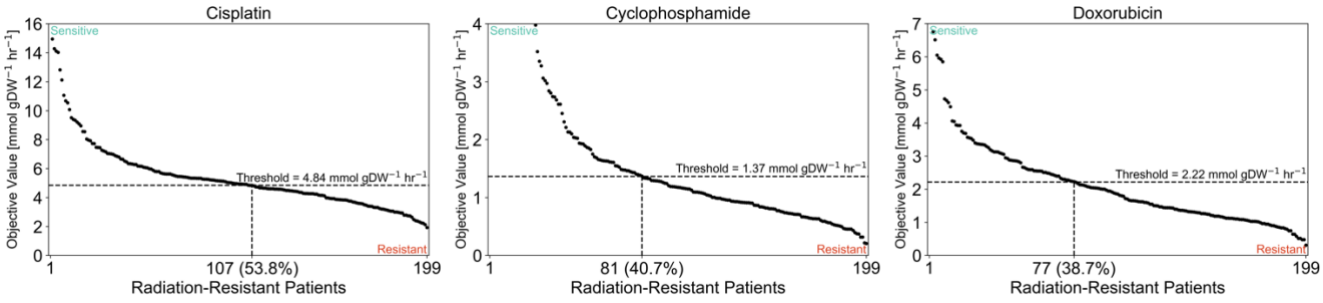


Figure 6-7. Prediction of chemotherapeutic response in radiation-resistant patient tumors. Optimal objective value thresholds separating drug-sensitive and -resistant classes are shown, with the number and percentage of radiation-resistant tumors predicted to be drug-sensitive.

While FBA models predicted that a large percentage of radiation-resistant patient tumors would not respond to treatment with these chemotherapies alone, it is yet unclear whether these chemotherapies could be used to sensitize these radiation-resistant tumors to radiation therapy when used in conjunction. The optimal objective function for implementation of cisplatin, cyclophosphamide, and doxorubicin metabolism involved the

redox cofactors GSH, NADH, and NADPH, respectively; all three of these reduced metabolites were previously shown to be upregulated in radiation-resistant tumor models for antioxidant protection against ROS-mediated damage from ionizing radiation (Figure 4-12). To predict the radiation-sensitizing effect of these chemotherapies in radiation-resistant drug-resistant patient tumors, the decrease in maximal GSH, NADH, and NADPH production caused by maintaining each tumor's cisplatin, cyclophosphamide, or doxorubicin objective value was determined (Figure 6-8). A wide spectrum of radiation-sensitizing effects was observed for each chemotherapy, with some tumor models showing significant decreases in reduced cofactor production, while others showing negligible effect.

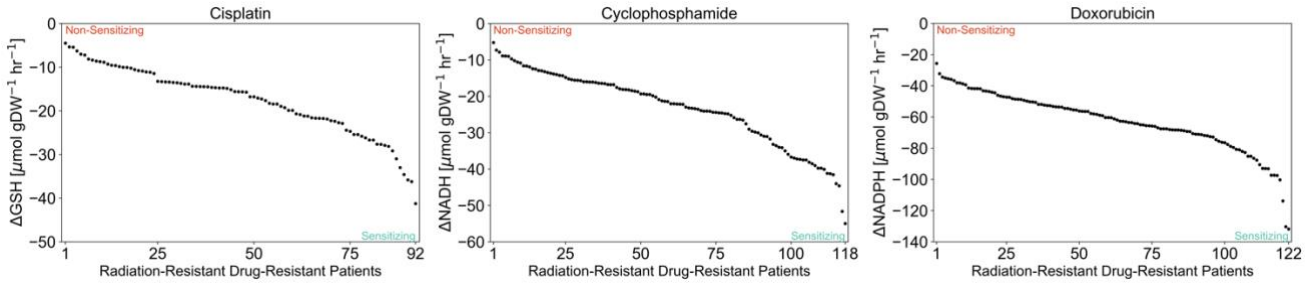


Figure 6-8. Predicted effect of maintaining chemotherapy objective value on maximal production of reduced redox cofactors in radiation-resistant drug-resistant tumor models.

6.3.3 Machine learning regressors identify sensitization biomarkers

To determine which multi-omic features are predictive of radiation-sensitizing effect and could be used as diagnostic biomarkers for combination therapy response, machine learning regressors utilizing clinical, gene expression, mutation, and metabolic features from each radiation-resistant drug-resistant TCGA tumor were built for each of

the three chemotherapies. Each machine learning regressor accurately predicts FBA model-calculated radiation-sensitizing effect of each chemotherapy, with statistically-significant Pearson and Spearman correlations between predicted and actual effects (Figure 6-9, Figure 6-10). The doxorubicin regressor shows the best performance among the three chemotherapies, likely due to the greater number of available radiation-resistant drug-resistant samples.

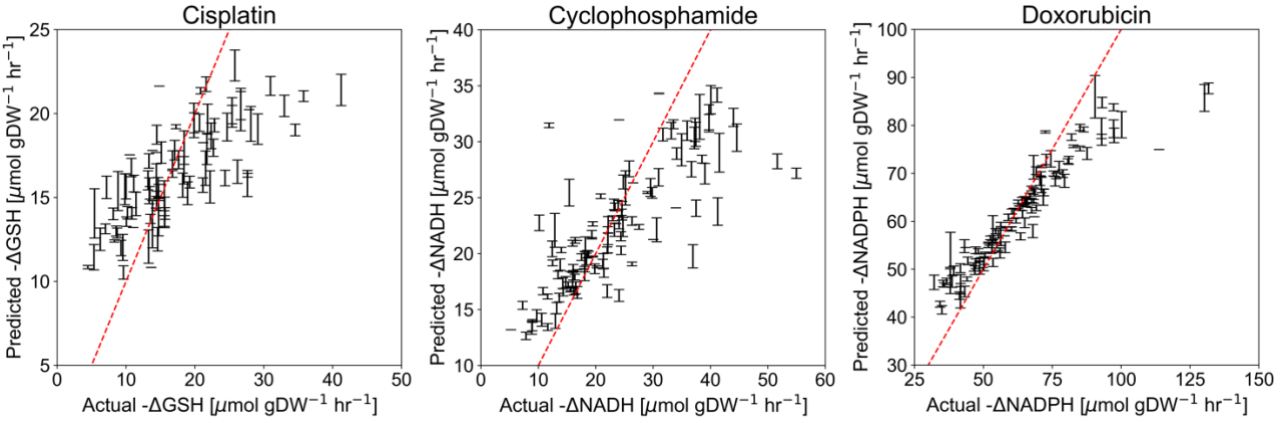


Figure 6-9. Comparison of machine learning regressor-predicted (y-axis) and FBA model-calculated (x-axis) radiation-sensitizing effects of each chemotherapy. Error

bars signify ± 1 standard error for each patient tumor across 20 training+validation/testing splits. Red dotted line: 1:1 line.

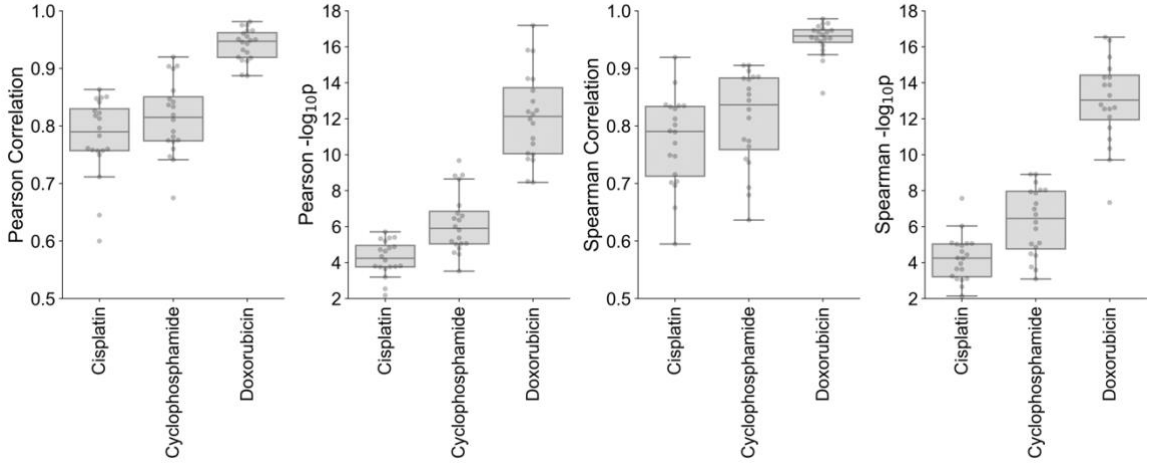


Figure 6-10. Pearson and Spearman correlation values (and associated p-values) for the regression of machine learning regression-predicted (y-axis) vs. FBA model-calculated (x-axis) radiation-sensitizing effects of each chemotherapy. All testing set performance values across 20 training+validation/testing splits are shown.

SHAP values were calculated for each feature among the three machine learning regressors to determine which multi-omic features are most predictive of radiation-sensitizing effect. Metabolic features contributed more than half of the cumulative feature importance scores for both cisplatin and cyclophosphamide regressors (Figure 6-11, Figure 6-12). Nicotinic acid adenine dinucleotide (NAAD), a 5,5'-dinucleotide metabolite that can be converted to NAD^+ via glutamine-dependent NAD^+ synthetase, was the 2nd-highest feature in the cisplatin regressor and highest overall feature in the cyclophosphamide regressor [393]. This metabolite is likely utilized in nicotinamide metabolism for generation of $NAD(P)^+$, which can be reduced to NAD(P)H and utilized by glutathione reductase to elevate levels of GSH for chemotherapy metabolism [394]. Additionally,

pantetheine 4'-phosphate, a related metabolite to pantothenic acid (vitamin B5) which is involved in lipid and cholesterol metabolism, as well as palmityl-CoA, a substrate for fatty acid elongation, were the highest and 3rd-highest features in the cisplatin regressor, respectively; cisplatin has been previously shown to cause hyperlipidemia in patients with germ cell testicular tumors, and sensitivity to cisplatin was shown to be affected by fatty acid synthase activity as well as administration of omega-3 fatty acids, validating the observed interplay between cisplatin and lipid metabolism [395-397]. $NADP^+$ and NADPH were the 4th- and 2nd-highest features in the cyclophosphamide regressor, respectively, suggesting that both NADH metabolism and NADPH metabolism are involved in the radiation-sensitizing effect of cyclophosphamide [117, 118, 398]. POR expression dominated the list of top features for the doxorubicin regressor, contributing nearly 45% of the cumulative feature importance scores (Figure 6-13). Interestingly, the POR NADPH-consuming reaction had a lower relative weight than the NOS3 NAPDH-consuming reaction in the optimal multi-feature objective function for doxorubicin metabolism; however, while NOS3 flux may better differentiate drug-sensitive and -resistant cancers, FBA models predict that the POR reaction has higher absolute flux values and causes greater reductions in NADPH stores, thus having a larger impact on radiation sensitization by doxorubicin (Figure 6-14).

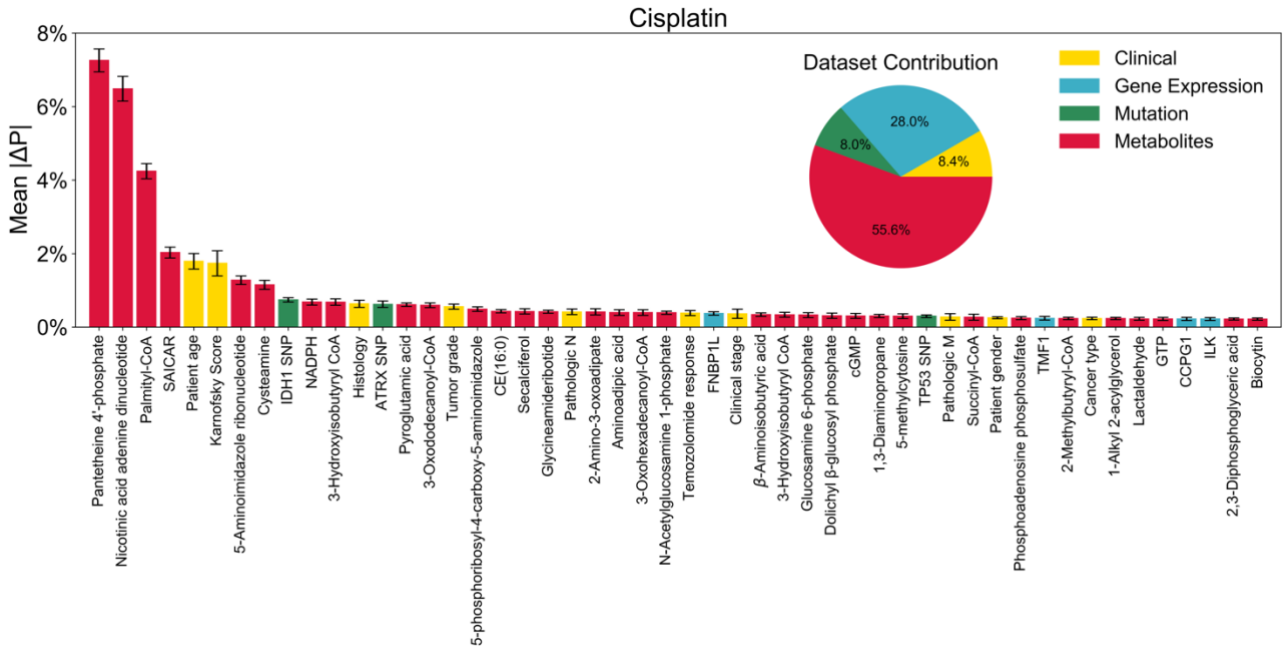


Figure 6-11. Top 50 features with largest mean absolute SHAP values for the cisplatin sensitizing-effect regressor, colored based on each feature's original dataset. (Inset) Relative contribution of features from each dataset to the sum of absolute SHAP values, averaged across all samples.

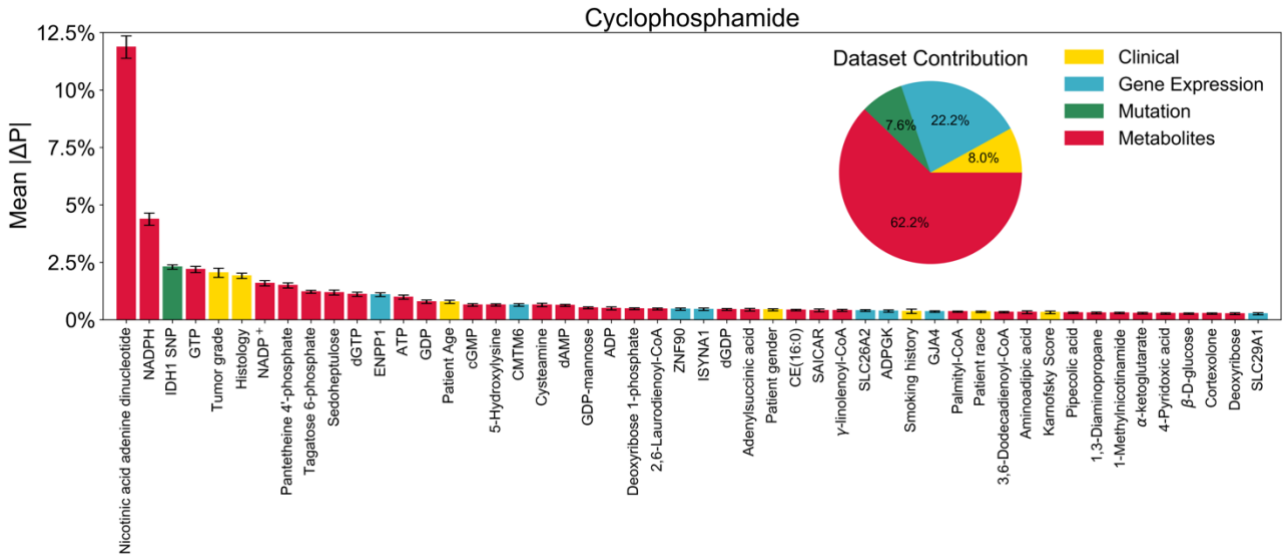


Figure 6-12. Top 50 features with largest mean absolute SHAP values for the cyclophosphamide sensitizing-effect regressor, colored based on each feature's original dataset. (Inset) Relative contribution of features from each dataset to the sum of absolute SHAP values, averaged across all samples.

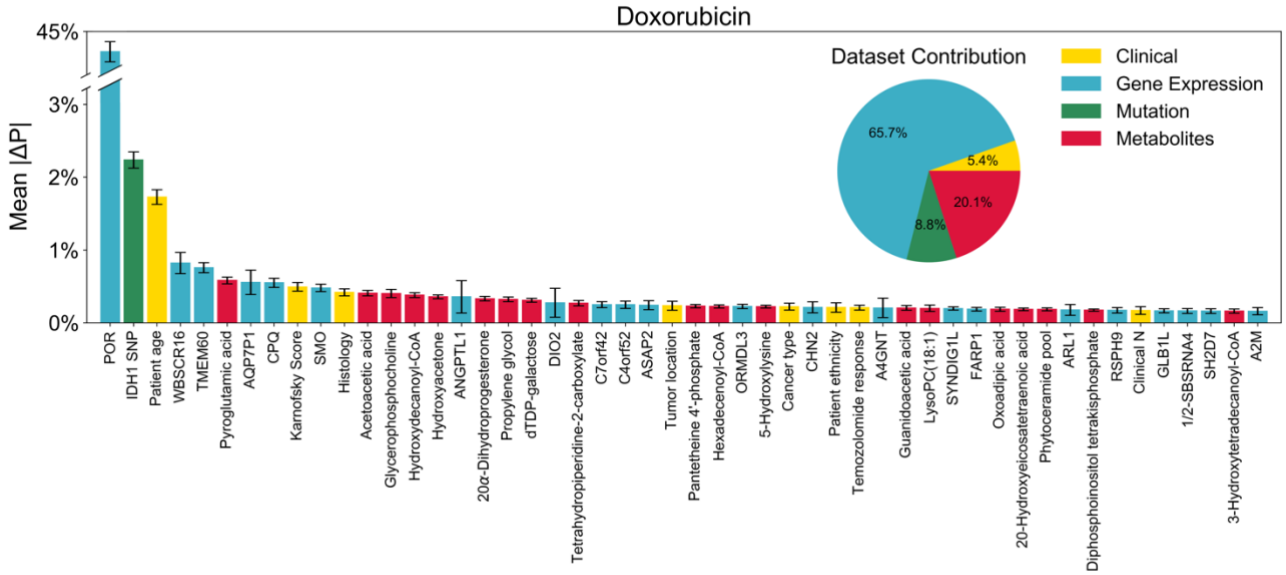


Figure 6-13. Top 50 features with largest mean absolute SHAP values for the doxorubicin sensitizing-effect regressor, colored based on each feature's original dataset. (Inset) Relative contribution of features from each dataset to the sum of absolute SHAP values, averaged across all samples.

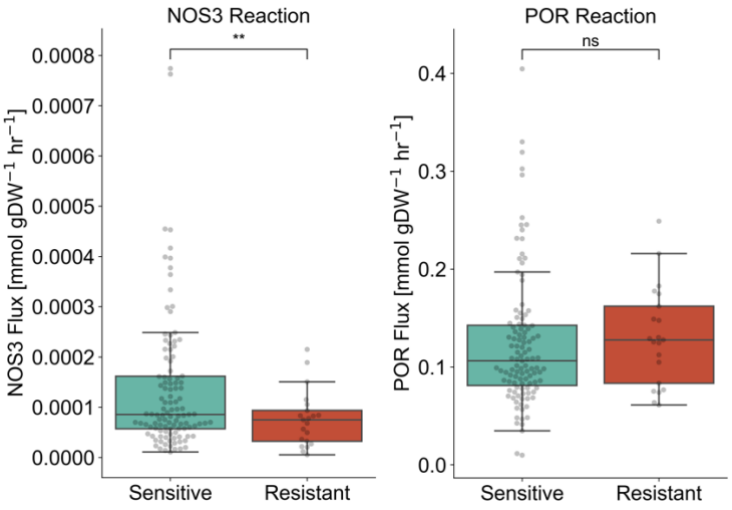


Figure 6-14. Comparison of NOS3 and POR flux magnitudes in doxorubicin-sensitive and -resistant tumor models.

6.4 Discussion

Because many chemotherapeutics with redox activity such as cisplatin, cyclophosphamide, and doxorubicin have multiple mechanisms of action, it is unclear which of these mechanisms contribute most to differential activity in drug-sensitive and -resistant tumors, as well as which confer the largest radiation-sensitizing effect. By implementing the mechanism of action of these three drugs into genome-scale FBA models of TCGA tumors and performing a novel multi-feature objective function screen, we found that maximization of multiple drug-related metabolic reactions or production of metabolic cofactors resulted in optimal correspondence between model predictions and known drug sensitivities in TCGA tumors, in agreement with the multiple observed mechanisms of action of these drugs (Figure 6-5). Feature weights provide an indication of the relative importance of different drug mechanisms to overall activity. Production of Cis-mtDNA had the highest feature weight in the cisplatin drug module, coinciding with experimental

studies demonstrating the significant impact of oxidative phosphorylation disruption on cisplatin-mediated oxidative stress (Figure 6-2) [106-108]. NADH production had the highest feature weight in the cyclophosphamide drug module; Magni et al. previously showed that overexpression of NADH-generating aldehyde dehydrogenase results in increased cyclophosphamide resistance in human hematopoietic cell lines, in agreement with model-predicted involvement of NADH in the major mechanism of cyclophosphamide activity (Figure 6-3) [118]. ATP production had the highest feature weight in the doxorubicin drug module, in agreement with previous studies implicating ATP-dependent ABC transporters in doxorubicin efflux and resistance (Figure 6-4) [399, 400]. Minor features with lower objective function weights also agree with previous experimental associations with drug sensitivity, including ATP-dependent efflux of cisplatin and glutathionylated cisplatin, as well as NADPH-consuming conversion of doxorubicin quinone to semiquinone form via NOS3 and POR [401, 402]. Interestingly, ATP production had the highest feature weight for doxorubicin but lowest feature weight for cisplatin, suggesting that the relative contribution of redox activity versus drug efflux towards overall chemotherapeutic response differs between drugs.

Implementation of optimal multi-feature objective functions for each drug module resulted in depletion of intracellular stores of GSH (cisplatin), NADH (cyclophosphamide), or NADPH (doxorubicin), all of which were upregulated in radiation-resistant FBA tumor models (Figure 4-12). Although redox-directed mechanisms of radiosensitization by these drugs have not been definitively identified, these three cofactors have been previously implicated in activity of the corresponding chemotherapies as well as overall chemotherapeutic sensitivity. Cisplatin-resistant cells have elevated levels of

GSH and increased expression of glutamate-cysteine ligase, the rate-limiting enzyme for GSH synthesis [403, 404]. Additionally, GSH depletion with L-BSO resulted in sensitization of HNSCC cell lines to cisplatin treatment, as well as partially inhibiting DNA damage repair in ovarian cancer cell lines [405, 406]. Thus, cisplatin and radiation response could be interconnected by cellular glutathione stores. Similarly, as mentioned above, elevated levels of NADH caused by increased ALDH expression are associated with cyclophosphamide resistance [118]. Doxorubicin treatment in yeast was found to cause increased expression of pentose phosphate pathway genes involved in NADPH generation, possibly suggesting an initial depletion in NADPH levels following treatment and compensatory transcriptional regulation [407, 408]. Inhibition of PGD resulted in increased doxorubicin sensitivity in thyroid cancer cells [409]. Taken together, these findings suggest that the redox cofactors predicted from FBA models are directly involved in the mechanism of action of these chemotherapeutics; nonetheless, further studies are needed to validate chemotherapy-induced depletion of redox cofactors and resultant radiosensitization.

Metabolism of these three chemotherapeutics involves the production and consumption of multiple metabolic cofactors including ATP, NADH, and NADPH. Most kinetic models of drug metabolism either introduce artificial supply reactions of these metabolites or the kinetics of only one of the major cofactor suppliers; for example, supply of NADPH is often introduced solely by G6PD. However, previous modeling of redox metabolism has shown the importance of multiple redox cofactor-producing reactions towards the overall metabolic phenotype, especially in radiation-resistant cancer cells (Figure 4-16) [206]. Because genome-scale metabolic models already include all known metabolic pathways involved in cofactor production and consumption, use of these models

provides an advantage over conventionally-used kinetic models by fully integrating drug metabolism into the non-xenobiotic human metabolic network [7]. This not only results in more accurate predictions of cofactor utilization in drug metabolism, but also allows for genome-scale screens for the identification of novel metabolic enzymes which most greatly impact chemotherapeutic response. Ultimately, integration of genome-scale steady-state models with small-scale kinetic models of xenobiotic metabolism may provide optimal performance in predicting cellular response to redox-based chemotherapeutic agents.

CHAPTER 7. CONCLUSIONS AND FUTURE DIRECTIONS

7.1 Conclusions

The overall goal of this dissertation was to make methodological advancements in the development and integration of multi-omic data into genome-scale metabolic models and machine learning classifiers, in order to accurately assess the role of redox metabolism in the tumor response to radiation therapy and radiation-sensitizing chemotherapies. By doing so, I was able to identify novel targets of redox metabolism in radiation-resistant tumors, predict radiation response in individual patient tumors through the use of metabolic biomarkers, and model the radiation-sensitizing effect of redox-based chemotherapeutics.

7.1.1 FBA models of redox metabolism in TCGA tumors

To create genome-scale models which accurately reflect the metabolism of radiation-sensitive and -resistant tumors, a novel bioinformatics platform was developed for integrating genomic, transcriptomic, kinetic, and thermodynamic parameters into personalized FBA models. By comparing model predictions from 716 radiation-sensitive and 199 radiation-resistant TCGA tumor models, I discovered that radiation-resistant tumors were capable of re-routing metabolic flux to optimize mitochondrial production of redox cofactors and increase ROS scavenging. Additionally, by performing simulated gene knockout screens and validating experimentally with siRNA gene knockdown, I identified metabolic genes (including G6PD, LDHB, PGAM2, and MTHFD1) which have significantly greater impacts on redox metabolism in radiation-resistant tumors and cancer cell lines across multiple cancer types. Finally, I took advantage of the personalized aspect of these metabolic models to generate metabolic flux profiles for individual radiation-

resistant patients, which can be used to identify which targets in the redox metabolic network may be optimal candidates for targeted therapies. This work not only improved upon the methodological shortcomings of previous FBA models of cancer metabolism, but was the first to utilize genome-scale modeling to identify key metabolic differences between radiation-sensitive and -resistant tumors which could be exploited for improving radiation sensitivity.

7.1.2 Machine learning classifiers for prediction of radiation response

To identify potential biomarkers for the *a priori* prediction of radiation sensitivity in cancer patients, machine learning classifiers which integrate multi-omic data from TCGA were developed. The predictive accuracy of multiple machine learning approaches was compared, identifying a dataset-independent ensemble architecture with gradient boosting models and Bayesian optimization for accurate prediction of radiation response compared to previously-developed classifiers. In order to incorporate metabolomic information into machine learning predictions, FBA models were used to identify metabolites which significant differences in production between radiation-sensitive and -resistant tumors; these metabolites were validated through experimental metabolomics studies of matched radiation-sensitive and -resistant cancer cell lines. Multi-omic and non-invasive classifiers identified clinical patient subgroups where metabolic features involved in lipid, nucleotide, antioxidant, and immune mediator metabolism provided significant diagnostic utility in predicting radiation therapy response. This work made significant advancements by being the first to integrate FBA model predictions into machine learning classifiers for cancer treatment outcomes, identifying key multi-omic biomarkers with clinical utility for assessment of radiation response in individual cancer patients.

7.1.3 FBA models of radiation-sensitizing chemotherapeutics

To determine which radiation-resistant patients could benefit from treatment with radiation-sensitizing chemotherapies, either alone or in combination with radiation therapy, the mechanisms of action of cisplatin, cyclophosphamide, and doxorubicin were incorporated into FBA models of radiation-resistant tumors. Optimal predictive accuracy of chemotherapeutic response was obtained by implementing a novel multi-feature FBA objective function screen, determining which metabolic reactions and cofactors are most important to the drugs' redox-based mechanisms of action. Model predicted-response from all three drugs showed decreases in levels of redox cofactors which had been previously implicated in tumor resistance to radiation therapy. Finally, machine learning regressors were utilized to identify multi-omic biomarkers from radiation-resistant patient tumors which are associated with redox cofactor depletion and radiation sensitization by these therapies. This work was the first to utilize genome-scale modeling to assess the role of chemotherapeutic treatment on tumor redox metabolism and radiation sensitization.

7.2 Future Directions

7.2.1 FBA models of redox metabolism in TCGA tumors

Although our bioinformatics and modeling approach towards predicting metabolic phenotypes in individual patient tumors using multi-omic data represents a significant methodological advancement over previous FBA models, additional improvements and integrations with other modeling strategies would further improve its accuracy and applicability. Currently, rates of metabolite transport through specific plasma membrane transporters are much less characterized than turnover rates of intracellular metabolic

enzymes, limiting the implementation of quantitative transport constraints [137]. Unless experimentally-measured metabolite uptake rates from samples of interest are obtained, constraints on uptake reactions are commonly set as binary (i.e. if a metabolite is present in the extracellular medium, intracellular uptake is unconstrained; otherwise, uptake is set to zero). A potential approach towards setting quantitative uptake constraints would be to relate membrane transporter expression with experimentally-measured extracellular metabolite concentrations (for example, from cell culture media or patient blood samples) to predict individualized metabolite uptake rates. In addition, regarding the application of FBA models towards studying redox metabolism, integration of important redox signaling and regulatory networks such as the Nrf2/Keap1/ARE pathway through methods such as integrated FBA (iFBA) could improve the accuracy of predicted reaction fluxes which utilize ARE-regulating genes including G6PD, PGD, IDH1, and ME1 [410-414]. Finally, the incorporation of FBA into agent-based tumor models using single-cell transcriptomic and mutational information may help explore the role of intratumor heterogeneity in redox metabolism and radiation responses [415, 416]. By simulating a multicellular system over time, the steady state metabolite import and export rates from one cell at a particular time point alter the import/export rates of these metabolites and resultant intracellular metabolism of neighboring cells at later time points; thus, dynamic metabolic behavior of the tumor and the impact of metabolic heterogeneity between cells can be assessed.

FBA models of radiation-sensitive and -resistant TCGA tumors could be used to explore other aspects of genome-scale metabolism which impact radiation response and could provide insight into optimal radiation-sensitizing therapies. Differences in energy metabolism and ATP production between radiation-sensitive and -resistant cancers have

been previously identified, including the promotion of radiation resistance by upregulation of aerobic glycolysis (Warburg effect) [179, 417]. For example, elevated lactate levels have been shown to reinforce DNA damage repair via inactivation of histone deacetylase [418]. FBA models could be used to identify key metabolic enzymes associated with aerobic glycolysis and lactate production in radiation-resistant cancers for targeted disruption of tumor energy metabolism. Additionally, previous computational studies utilizing genome-scale modeling have explored the role of dietary changes on metabolism and disease progression [419]. Metabolite import rates in personalized FBA models can be used to reflect a particular cancer patient's diet, and the resulting impact on genome-scale tumor metabolism can be explored to identify optimal nutrition strategies for individual patients. The flexibility and genome-scale coverage of FBA models enables their utilizing in myriad applications for predicting metabolic phenotypes and discovering metabolic targets for improving treatment response.

7.2.2 Machine learning classifiers for prediction of radiation response

Although our novel approach towards integrating machine learning and genome-scale metabolic modeling for the prediction of radiation response provides many enhancements in performance and biological insights compared to previous studies, additional improvements could be made. Analysis of the multi-omic classifier showed that tumor histology had the largest average SHAP value across all datasets (Figure 5-26). This suggests that additional features from tumor histological images could further improve the predictive accuracy of classifiers for radiation response. A convolutional neural network could be added as an additional base learner, with tumor H&E images from TCGA provided as input, to both enhance classifier predictions and identify novel connections

between tumor histological features and radiation sensitivity [420, 421]. Additionally, changes in both DNA methylation and microRNA expression have been implicated in the tumor response to ionizing radiation therapy and may alter overall patient sensitivity to therapy [422-425]. Inclusion of these features into the multi-omic classifier may yield additional multi-omic biomarkers for improved radiation response prediction. To ensure that the obtained pan-cancer classifiers and diagnostic biomarkers for radiation response are robust and not cancer type-dependent, an adversarial penalty could be added to the weighted log loss function to punish classifiers which discriminate based on the cancer type covariate [426]. Finally, because the TCGA database is biased towards patients with advanced disease compared to the general cancer patient population, the tumor pathophysiology and thus the multi-omic datasets from these tumor samples may not be reflective of early-stage disease patients, limiting the generalizability of the classifiers developed from these datasets. The inclusion of samples from another tumor database as either interspersed with TCGA samples or as a separate testing dataset could help improve the predictive accuracy and generalizability of these classifiers.

While this study focused on the identification of diagnostic biomarkers for predicting response to radiation therapy, the integration of machine learning and metabolic models is a generalizable approach that could be used for many other applications in cancer biology and treatment. For example, kinetic computational models are widely used for studying chemotherapeutic metabolism, and genome-scale metabolic models have recently been adapted for the prediction of chemotherapeutic response [383, 427]. These tools could be used to obtain metabolic predictions from drug-sensitive and -resistant tumors, which can be integrated with other multi-omic datasets for identifying biomarkers for drug

response. Flux variance analysis could also be used to obtain individual reaction flux values across the human metabolic network for each patient tumor; integration into machine learning classifiers would then identify the metabolic reactions with activities most associated with treatment sensitivity or resistance, providing useful insights into cancer metabolism at a systems level. Finally, multiple types of model outputs, including objective values, reaction fluxes, and gene knockout predictions, can be integrated into machine learning classifiers to determine which types of FBA analysis provide the greatest utility in predicting metabolic phenotypes and should be given precedence in terms of computational resources and experimental validation.

7.2.3 FBA models of radiation-sensitizing chemotherapeutics

While FBA models of cisplatin, cyclophosphamide, and doxorubin metabolism suggested that the depletion of redox cofactors may contribute to the overall radiation-sensitizing effects of these chemotherapies, further experimental studies are needed to validate these effects. After the depletion of redox cofactors following chemotherapeutic treatment is verified, pharmacologic inhibition of cofactor production (e.g. L-BSO treatment to prevent glutathione production) or siRNA knockdown of cofactor-producing genes (e.g. NADK to prevent $NADP_+$ or NADPH production) can be used in radiation-resistant cell lines to assess whether increased sensitivity to radiation treatment occurs. Subsequently, the contribution of redox cofactor depletion towards the overall radiosensitizing effect of these chemotherapies can be established by comparing cell viability following drug treatment with or without exogenous addition of the redox cofactors that are being depleted. Finally, the involvement of specific metabolic reactions identified from optimal multi-feature objective functions and the relative weights between

features can be validated through pharmacologic inhibition or siRNA knockdown of the associated enzyme-encoding genes.

Integration of genome-scale modeling with kinetic models of xenobiotic metabolism can be used to combine time-resolved drug pharmacodynamics with network-driven cofactor supplies for increased predictive accuracy. Similar integrative approaches for metabolic modeling have been developed previously for microbiological applications [428]. Feng et al. combined a genome-scale model of the bacterium *Shewanella oneidensis* MR-1 with kinetic-based bioprocess models to capture transient alterations in intracellular flux distributions with different growth stages [429]. Chowdhury et al. developed k-OptForce for combining FBA with kinetic models incorporating substrate-level enzyme regulation to identify optimal network-wide intervention strategies for enhanced biochemical production [430]. Because both chemotherapeutic activity and redox biology rely on metabolic processes at short time-scales (ROS production, cofactor utilization) as well as longer time-scales (drug clearance, induction of gene expression changes), integrative modeling approaches may be necessary for accurate prediction of redox-based chemotherapeutic response and radiation sensitization in radiation-resistant tumors.

REFERENCES

1. Miller, K.D., et al., *Cancer treatment and survivorship statistics, 2016.* CA Cancer J Clin, 2016. **66**(4): p. 271-89.
2. Weinstein, J.N., et al., *The Cancer Genome Atlas Pan-Cancer analysis project.* Nat Genet, 2013. **45**(10): p. 1113-20.
3. Skvortsov, S., et al., *Putative biomarkers and therapeutic targets associated with radiation resistance.* Expert Rev Proteomics, 2014. **11**(2): p. 207-14.
4. Ward, J.F., et al., *Radiation and hydrogen peroxide induced free radical damage to DNA.* Br J Cancer Suppl, 1987. **8**: p. 105-12.
5. Sun, J., et al., *Role of antioxidant enzymes on ionizing radiation resistance.* Free Radic Biol Med, 1998. **24**(4): p. 586-93.
6. Deavall, D.G., et al., *Drug-induced oxidative stress and toxicity.* Journal of toxicology, 2012. **2012**.
7. Brunk, E., et al., *Recon3D enables a three-dimensional view of gene variation in human metabolism.* Nat Biotechnol, 2018. **36**(3): p. 272-281.
8. Kang, J., et al., *Radiogenomics models in precision radiotherapy: from mechanistic to machine learning.* arXiv preprint arXiv:1904.09662, 2019.
9. Orth, J.D., I. Thiele, and B.O. Palsson, *What is flux balance analysis?* Nat Biotechnol, 2010. **28**(3): p. 245-8.
10. Nilsson, A. and J. Nielsen, *Genome scale metabolic modeling of cancer.* Metab Eng, 2017. **43**(Pt B): p. 103-112.
11. Blazier, A.S. and J.A. Papin, *Integration of expression data in genome-scale metabolic network reconstructions.* Front Physiol, 2012. **3**: p. 299.
12. Henry, C.S., L.J. Broadbelt, and V. Hatzimanikatis, *Thermodynamics-based metabolic flux analysis.* Biophys J, 2007. **92**(5): p. 1792-805.
13. Delaney, G., et al., *The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines.* Cancer, 2005. **104**(6): p. 1129-37.
14. Baskar, R., et al., *Cancer and radiation therapy: current advances and future directions.* Int J Med Sci, 2012. **9**(3): p. 193-9.
15. Kim, B.M., et al., *Therapeutic Implications for Overcoming Radiation Resistance in Cancer Therapy.* Int J Mol Sci, 2015. **16**(11): p. 26880-913.
16. Tang, L., et al., *Role of metabolism in cancer cell radioresistance and radiosensitization methods.* J Exp Clin Cancer Res, 2018. **37**(1): p. 87.
17. Baskar, R., et al., *Biological response of cancer cells to radiation treatment.* Front Mol Biosci, 2014. **1**: p. 24.
18. Barcellos-Hoff, M.H., C. Park, and E.G. Wright, *Radiation and the microenvironment - tumorigenesis and therapy.* Nat Rev Cancer, 2005. **5**(11): p. 867-75.
19. Jackson, S.P., *Sensing and repairing DNA double-strand breaks.* Carcinogenesis, 2002. **23**(5): p. 687-96.

20. Luo, Y., E.S. Henle, and S. Linn, *Oxidative damage to DNA constituents by iron-mediated fenton reactions. The deoxycytidine family.* J Biol Chem, 1996. **271**(35): p. 21167-76.
21. Shaw, P.H., *The role of p53 in cell cycle regulation.* Pathol Res Pract, 1996. **192**(7): p. 669-75.
22. Chen, J., *The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression.* Cold Spring Harb Perspect Med, 2016. **6**(3): p. a026104.
23. Blackford, A.N. and S.P. Jackson, *ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response.* Mol Cell, 2017. **66**(6): p. 801-817.
24. Yang, J., et al., *ATM, ATR and DNA-PK: initiators of the cellular genotoxic stress responses.* Carcinogenesis, 2003. **24**(10): p. 1571-80.
25. Maier, P., et al., *Cellular Pathways in Response to Ionizing Radiation and Their Targetability for Tumor Radiosensitization.* Int J Mol Sci, 2016. **17**(1).
26. Sabin, R.J. and R.M. Anderson, *Cellular Senescence - its role in cancer and the response to ionizing radiation.* Genome Integr, 2011. **2**(1): p. 7.
27. Surova, O. and B. Zhivotovsky, *Various modes of cell death induced by DNA damage.* Oncogene, 2013. **32**(33): p. 3789-97.
28. Vogin, G. and N. Foray, *The law of Bergonie and Tribondeau: a nice formula for a first approximation.* Int J Radiat Biol, 2013. **89**(1): p. 2-8.
29. Bergonie, J., *De quelques resultas de la radiotherapie et essai de fixation d'une technique radionnelle.* CR Acad. Sci., 1906. **143**: p. 983-995.
30. Kimbrough, R.A. and C.C. Norris, *Factors influencing end-results in carcinoma of the cervix after irradiation.* American Journal of Obstetrics & Gynecology, 1927. **13**(3): p. 279-287.
31. Kistner, R.W. and A.T. Hertig, *A correlation of histologic grade, clinical state, and radiation response in carcinoma of the uterine cervix.* Am J Obstet Gynecol, 1951. **61**(6): p. 1293-1300.
32. Kagan, A.R., et al., *The success of treatment of brain metastasis with radiation therapy.* International Journal of Radiation Oncology• Biology• Physics, 1977. **2**: p. 74-75.
33. Griffin, T.W., et al., *Predicting the response of head and neck cancers to radiation therapy with a multivariate modelling system: an analysis of the RTOG head and neck registry.* Int J Radiat Oncol Biol Phys, 1984. **10**(4): p. 481-7.
34. Mottram, J., *A factor of importance in the radio sensitivity of tumours.* The British Journal of Radiology, 1936. **9**(105): p. 606-614.
35. Nordsmark, M., M. Overgaard, and J. Overgaard, *Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head and neck.* Radiother Oncol, 1996. **41**(1): p. 31-9.
36. Fyles, A.W., et al., *Oxygenation predicts radiation response and survival in patients with cervix cancer.* Radiother Oncol, 1998. **48**(2): p. 149-56.
37. Foray, N., C. Colin, and M. Bourguignon, *100 years of individual radiosensitivity: how we have forgotten the evidence.* Radiology, 2012. **264**(3): p. 627-31.

38. Foray, N., et al., *Hypersensitivity of ataxia telangiectasia fibroblasts to ionizing radiation is associated with a repair deficiency of DNA double-strand breaks.* Int J Radiat Biol, 1997. **72**(3): p. 271-83.
39. Wang, M., et al., *Radiation Resistance in KRAS-Mutated Lung Cancer Is Enabled by Stem-like Properties Mediated by an Osteopontin-EGFR Pathway.* Cancer Res, 2017. **77**(8): p. 2018-2028.
40. Farquhar, D., et al., *Genetic Mutations in KEAP/NFE2L2 Associated With Radiation Resistance in Early-Stage Laryngeal Squamous Cell Carcinoma: A Case Series.* International Journal of Radiation Oncology• Biology• Physics, 2018. **100**(5): p. 1375-1376.
41. Choudhury, A., et al., *MRE11 expression is predictive of cause-specific survival following radical radiotherapy for muscle-invasive bladder cancer.* Cancer Res, 2010. **70**(18): p. 7017-26.
42. Torres-Roca, J.F., *A molecular assay of tumor radiosensitivity: a roadmap towards biology-based personalized radiation therapy.* Per Med, 2012. **9**(5): p. 547-557.
43. Eschrich, S.A., et al., *Validation of a radiosensitivity molecular signature in breast cancer.* Clin Cancer Res, 2012. **18**(18): p. 5134-43.
44. Creelan, B., et al., *A gene expression platform to predict benefit from adjuvant external beam radiation in resected non-small cell lung cancer.* International Journal of Radiation Oncology• Biology• Physics, 2014. **90**(1): p. S76-S77.
45. Torres-Roca, J., et al., *A molecular signature of radiosensitivity (RSI) is an RT-specific biomarker in prostate cancer.* International Journal of Radiation Oncology• Biology• Physics, 2014. **90**(1): p. S157.
46. Manem, V.S. and A. Dhawan, *RadiationGeneSigDB: a database of oxic and hypoxic radiation response gene signatures and their utility in pre-clinical research.* Br J Radiol, 2019. **92**(1103): p. 20190198.
47. Verdin, E., *NAD+ in aging, metabolism, and neurodegeneration.* Science, 2015. **350**(6265): p. 1208-1213.
48. Garrido, A. and N. Djouder, *NAD+ deficits in age-related diseases and cancer.* Trends in cancer, 2017. **3**(8): p. 593-610.
49. Kennedy, B.E., et al., *NAD+ salvage pathway in cancer metabolism and therapy.* Pharmacological research, 2016. **114**: p. 274-283.
50. Michels, J., et al., *PARP and other prospective targets for poisoning cancer cell metabolism.* Biochemical pharmacology, 2014. **92**(1): p. 164-171.
51. Xiao, W., et al., *NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism.* Antioxid Redox Signal, 2018. **28**(3): p. 251-272.
52. Houtkooper, R.H., et al., *The secret life of NAD+: an old metabolite controlling new metabolic signaling pathways.* Endocrine reviews, 2010. **31**(2): p. 194-223.
53. Revollo, J.R., A.A. Grimm, and S.-i. Imai, *The NAD biosynthesis pathway mediated by nicotinamide phosphoribosyltransferase regulates Sir2 activity in mammalian cells.* Journal of Biological Chemistry, 2004. **279**(49): p. 50754-50763.
54. Schafer, F.Q. and G.R. Buettner, *Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple.* Free radical biology and medicine, 2001. **30**(11): p. 1191-1212.

55. Lin, S.-J. and L. Guarente, *Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease.* Current opinion in cell biology, 2003. **15**(2): p. 241-246.
56. Sies, H., *Metabolic compartmentation*. 1982: Academic Pr.
57. Franklin, D.A., et al., *p53 coordinates DNA repair with nucleotide synthesis by suppressing PFKFB3 expression and promoting the pentose phosphate pathway.* Sci Rep, 2016. **6**: p. 38067.
58. Lewis, C.A., et al., *Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells.* Mol Cell, 2014. **55**(2): p. 253-63.
59. Hoxhaj, G., et al., *Direct stimulation of NADP(+) synthesis through Akt-mediated phosphorylation of NAD kinase.* Science, 2019. **363**(6431): p. 1088-1092.
60. Stanton, R.C., *Glucose-6-phosphate dehydrogenase, NADPH, and cell survival.* IUBMB Life, 2012. **64**(5): p. 362-9.
61. Cho, E.S., et al., *The Pentose Phosphate Pathway as a Potential Target for Cancer Therapy.* Biomol Ther (Seoul), 2018. **26**(1): p. 29-38.
62. Zitka, O., et al., *Redox status expressed as GSH:GSSG ratio as a marker for oxidative stress in paediatric tumour patients.* Oncol Lett, 2012. **4**(6): p. 1247-1253.
63. Kranner, I., et al., *Glutathione half-cell reduction potential: a universal stress marker and modulator of programmed cell death?* Free Radic Biol Med, 2006. **40**(12): p. 2155-65.
64. Lu, S.C., *Glutathione synthesis.* Biochim Biophys Acta, 2013. **1830**(5): p. 3143-53.
65. Yu, X. and Y.C. Long, *Crosstalk between cystine and glutathione is critical for the regulation of amino acid signaling pathways and ferroptosis.* Sci Rep, 2016. **6**: p. 30033.
66. Fukai, T. and M. Ushio-Fukai, *Superoxide dismutases: role in redox signaling, vascular function, and diseases.* Antioxid Redox Signal, 2011. **15**(6): p. 1583-606.
67. Chance, B., *Effect of pH upon the reaction kinetics of the enzyme-substrate compounds of catalase.* J Biol Chem, 1952. **194**(2): p. 471-81.
68. Purdue, P.E. and P.B. Lazarow, *Targeting of human catalase to peroxisomes is dependent upon a novel COOH-terminal peroxisomal targeting sequence.* J Cell Biol, 1996. **134**(4): p. 849-62.
69. Kirkman, H.N., et al., *Mechanisms of protection of catalase by NADPH. Kinetics and stoichiometry.* J Biol Chem, 1999. **274**(20): p. 13908-14.
70. Dey, S., A. Sidor, and B. O'Rourke, *Compartment-specific Control of Reactive Oxygen Species Scavenging by Antioxidant Pathway Enzymes.* J Biol Chem, 2016. **291**(21): p. 11185-97.
71. Morgan, M.A. and T.S. Lawrence, *Molecular Pathways: Overcoming Radiation Resistance by Targeting DNA Damage Response Pathways.* Clin Cancer Res, 2015. **21**(13): p. 2898-904.
72. Santivasi, W.L. and F. Xia, *Ionizing radiation-induced DNA damage, response, and repair.* Antioxid Redox Signal, 2014. **21**(2): p. 251-9.
73. Hileman, E.A., G. Achanta, and P. Huang, *Superoxide dismutase: an emerging target for cancer therapeutics.* Expert Opin Ther Targets, 2001. **5**(6): p. 697-710.

74. Bauer, G., *Tumor cell-protective catalase as a novel target for rational therapeutic approaches based on specific intercellular ROS signaling.* Anticancer Res, 2012. **32**(7): p. 2599-624.
75. Jiao, Y., et al., *Glutathione peroxidases as oncotargets.* Oncotarget, 2017. **8**(45): p. 80093-80102.
76. Marklund, S.L., et al., *Radiation resistance and the CuZn superoxide dismutase, Mn superoxide dismutase, catalase, and glutathione peroxidase activities of seven human cell lines.* Radiat Res, 1984. **100**(1): p. 115-23.
77. Alano, C.C., et al., *NAD+ depletion is necessary and sufficient forPoly (ADP-ribose) polymerase-1-mediated neuronal death.* Journal of Neuroscience, 2010. **30**(8): p. 2967-2978.
78. Moroni, F., *Poly (ADP-ribose) polymerase 1 (PARP-1) and postischemic brain damage.* Current opinion in pharmacology, 2008. **8**(1): p. 96-103.
79. Gujar, A.D., et al., *An NAD+-dependent transcriptional program governs self-renewal and radiation resistance in glioblastoma.* Proc Natl Acad Sci U S A, 2016. **113**(51): p. E8247-e8256.
80. Tran, A.N., et al., *Increased sensitivity to radiochemotherapy in IDH1 mutant glioblastoma as demonstrated by serial quantitative MR volumetry.* Neuro Oncol, 2014. **16**(3): p. 414-20.
81. Gelman, S.J., et al., *Consumption of NADPH for 2-HG Synthesis Increases Pentose Phosphate Pathway Flux and Sensitizes Cells to Oxidative Stress.* Cell Rep, 2018. **22**(2): p. 512-522.
82. Yang, H.C., et al., *The Redox Role of G6PD in Cell Growth, Cell Death, and Cancer.* Cells, 2019. **8**(9).
83. Catanzaro, D., et al., *Inhibition of glucose-6-phosphate dehydrogenase sensitizes cisplatin-resistant cells to death.* Oncotarget, 2015. **6**(30): p. 30102-14.
84. Li, D., et al., *Inhibition of glutamine metabolism counteracts pancreatic cancer stem cell features and sensitizes cells to radiotherapy.* Oncotarget, 2015. **6**(31): p. 31151-31163.
85. Fu, S., et al., *Glutamine Synthetase Promotes Radiation Resistance via Facilitating Nucleotide Metabolism and Subsequent DNA Damage Repair.* Cell Rep, 2019. **28**(5): p. 1136-1143.e4.
86. Heidelberger, C., et al., *Studies on fluorinated pyrimidines. II. Effects on transplanted tumors.* Cancer Res, 1958. **18**(3): p. 305-17.
87. Nigro, N.D., V.K. Vaitkevicius, and B. Considine, Jr., *Combined therapy for cancer of the anal canal: a preliminary report.* Dis Colon Rectum, 1974. **17**(3): p. 354-6.
88. Seiwert, T.Y., J.K. Salama, and E.E. Vokes, *The chemoradiation paradigm in head and neck cancer.* Nature clinical practice Oncology, 2007. **4**(3): p. 156-171.
89. Milano, M.T., et al., *Patterns and timing of recurrence after temozolomide-based chemoradiation for glioblastoma.* International Journal of Radiation Oncology* Biology* Physics, 2010. **78**(4): p. 1147-1155.
90. Gordon, G.S. and E.E. Vokes, *Chemoradiation for locally advanced, unresectable NSCLC.* Lung Cancer, 1999. **13**(8).

91. Stahl, M., et al., *Chemoradiation with and without surgery in patients with locally advanced squamous cell carcinoma of the esophagus.* Journal of clinical oncology, 2005. **23**(10): p. 2310-2317.
92. Kalady, M.F., et al., *Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer.* Annals of surgery, 2009. **250**(4): p. 582-589.
93. Lohrisch, C., et al., *Small cell carcinoma of the bladder: long term outcome with integrated chemoradiation.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 1999. **86**(11): p. 2346-2352.
94. Wong, L., et al., *Chemoradiation and adjuvant chemotherapy in cervical cancer.* Journal of clinical oncology, 1999. **17**(7): p. 2055-2055.
95. Xiong, H.Q., et al., *Chemoradiation for resectable gastric cancer.* The lancet oncology, 2003. **4**(8): p. 498-505.
96. Thoms, J. and R.G. Bristow, *DNA repair targeting and radiotherapy: a focus on the therapeutic ratio.* Semin Radiat Oncol, 2010. **20**(4): p. 217-22.
97. Calsou, P. and B. Salles, *Role of DNA repair in the mechanisms of cell resistance to alkylating agents and cisplatin.* Cancer chemotherapy and pharmacology, 1993. **32**(2): p. 85-89.
98. Ludeman, S.M., *The chemistry of the metabolites of cyclophosphamide.* Current pharmaceutical design, 1999. **5**: p. 627-644.
99. Bodley, A., et al., *DNA topoisomerase II-mediated interaction of doxorubicin and daunorubicin congeners with DNA.* Cancer research, 1989. **49**(21): p. 5969-5978.
100. Mitchell, J.B., et al., *In vitro and in vivo radiation sensitization of human tumor cells by a novel checkpoint kinase inhibitor, AZD7762.* Clin Cancer Res, 2010. **16**(7): p. 2076-84.
101. Huang, S.M. and P.M. Harari, *Modulation of radiation response after epidermal growth factor receptor blockade in squamous cell carcinomas: inhibition of damage repair, cell cycle kinetics, and tumor angiogenesis.* Clin Cancer Res, 2000. **6**(6): p. 2166-74.
102. Kvols, L.K., *Radiation sensitizers: a selective review of molecules targeting DNA and non-DNA targets.* Journal of Nuclear Medicine, 2005. **46**(1 suppl): p. 187S-190S.
103. Rosenberg, A. and S. Knox, *Radiation sensitization with redox modulators: a promising approach.* Int J Radiat Oncol Biol Phys, 2006. **64**(2): p. 343-54.
104. Rischin, D., et al., *Tirapazamine, cisplatin, and radiation versus fluorouracil, cisplatin, and radiation in patients with locally advanced head and neck cancer: a randomized phase II trial of the Trans-Tasman Radiation Oncology Group (TROG 98.02).* Journal of Clinical Oncology, 2005. **23**(1): p. 79-87.
105. Shaw, E., et al., *RSR13 plus cranial radiation therapy in patients with brain metastases: comparison with the Radiation Therapy Oncology Group Recursive Partitioning Analysis Brain Metastases Database.* Journal of clinical oncology, 2003. **21**(12): p. 2364-2371.
106. Marullo, R., et al., *Cisplatin induces a mitochondrial-ROS response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions.* PLoS One, 2013. **8**(11): p. e81162.

107. Masuda, H., et al., *Detection and cytotoxicity of cisplatin-induced superoxide anion in monolayer cultures of a human ovarian cancer cell line.* Cancer Chemother Pharmacol, 2001. **47**(2): p. 155-60.
108. Yang, Z., et al., *Cisplatin preferentially binds mitochondrial DNA and voltage-dependent anion channel protein in the mitochondrial membrane of head and neck squamous cell carcinoma: possible role in apoptosis.* Clin Cancer Res, 2006. **12**(19): p. 5817-25.
109. Akerfeldt, M.C., et al., *Interactions of cisplatin and the copper transporter CTR1 in human colon cancer cells.* J Biol Inorg Chem, 2017. **22**(5): p. 765-774.
110. Bharthuar, A., et al., *Breast cancer resistance protein (BCRP) and excision repair cross complement-1 (ERCC1) expression in esophageal cancers and response to cisplatin and irinotecan based chemotherapy.* J Gastrointest Oncol, 2014. **5**(4): p. 253-8.
111. Li, Y.Q., et al., *Copper efflux transporters ATP7A and ATP7B: Novel biomarkers for platinum drug resistance and targets for therapy.* IUBMB life, 2018. **70**(3): p. 183-191.
112. Zhu, S., et al., *A Role for The ATP7A Copper Transporter in Tumorigenesis and Cisplatin Resistance.* J Cancer, 2017. **8**(11): p. 1952-1958.
113. Zhao, L., et al., *Cisplatin binds to human copper chaperone Cox17: the mechanistic implication of drug delivery to mitochondria.* Chem Commun (Camb), 2014. **50**(20): p. 2667-9.
114. Smith, D.J., et al., *Metallothioneins and resistance to cisplatin and radiation in prostate cancer.* Urology, 2006. **67**(6): p. 1341-7.
115. De Luca, A., et al., *A structure-based mechanism of cisplatin resistance mediated by glutathione transferase P1-1.* Proceedings of the National Academy of Sciences, 2019. **116**(28): p. 13943-13951.
116. Hall, A. and M. Tilby, *Mechanisms of action of, and modes of resistance to, alkylating agents used in the treatment of haematological malignancies.* Blood reviews, 1992. **6**(3): p. 163-173.
117. Parekh, H.K. and N.E. Sladek, *NADPH-dependent enzyme-catalyzed reduction of aldophosphamide, the pivotal metabolite of cyclophosphamide.* Biochem Pharmacol, 1993. **46**(6): p. 1043-52.
118. Magni, M., et al., *Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer.* Blood, 1996. **87**(3): p. 1097-103.
119. Canitrot, Y., et al., *Multidrug resistance genes (MRP) and MDR1 expression in small cell lung cancer xenografts: relationship with response to chemotherapy.* Cancer Lett, 1998. **130**(1-2): p. 133-41.
120. Brayboy, L.M., et al., *Multidrug resistance transporter-1 and breast cancer resistance protein protect against ovarian toxicity, and are essential in ovarian physiology.* Reprod Toxicol, 2017. **69**: p. 121-131.
121. Cox, P.J., B.J. Phillips, and P. Thomas, *The enzymatic basis of the selective action of cyclophosphamide.* Cancer Res, 1975. **35**(12): p. 3755-61.
122. Gurtoo, H.L., J.H. Hipkens, and S.D. Sharma, *Role of glutathione in the metabolism-dependent toxicity and chemotherapy of cyclophosphamide.* Cancer Res, 1981. **41**(9 Pt 1): p. 3584-91.

123. Zhu, H., et al., *Doxorubicin Redox Biology: Redox Cycling, Topoisomerase Inhibition, and Oxidative Stress.* React Oxyg Species (Apex), 2016. **1**(3): p. 189-198.
124. Cappetta, D., et al., *Oxidative stress and cellular response to doxorubicin: a common factor in the complex milieu of anthracycline cardiotoxicity.* Oxidative medicine and cellular longevity, 2017. **2017**.
125. Finn, N.A., H.W. Findley, and M.L. Kemp, *A switching mechanism in doxorubicin bioactivation can be exploited to control doxorubicin toxicity.* PLoS Comput Biol, 2011. **7**(9): p. e1002151.
126. Okabe, M., et al., *Characterization of the organic cation transporter SLC22A16: a doxorubicin importer.* Biochem Biophys Res Commun, 2005. **333**(3): p. 754-62.
127. Singhal, S.S., et al., *Doxorubicin transport by RALBP1 and ABCG2 in lung and breast cancer.* Int J Oncol, 2007. **30**(3): p. 717-25.
128. Neilan, T.G., et al., *Disruption of nitric oxide synthase 3 protects against the cardiac injury, dysfunction, and mortality induced by doxorubicin.* Circulation, 2007. **116**(5): p. 506-14.
129. Mobaraki, M., et al., *Molecular mechanisms of cardiotoxicity: a review on major side-effect of doxorubicin.* Indian Journal of Pharmaceutical Sciences, 2017. **79**(3): p. 335-344.
130. Riddick, D.S., et al., *NADPH-cytochrome P450 oxidoreductase: roles in physiology, pharmacology, and toxicology.* Drug Metab Dispos, 2013. **41**(1): p. 12-23.
131. Gille, L. and H. Nohl, *Analyses of the molecular mechanism of adriamycin-induced cardiotoxicity.* Free Radical Biology and Medicine, 1997. **23**(5): p. 775-782.
132. Lewis, N.E., et al., *Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models.* Mol Syst Biol, 2010. **6**: p. 390.
133. Megchelenbrink, W., M. Huynen, and E. Marchiori, *optGpSampler: an improved tool for uniformly sampling the solution-space of genome-scale metabolic networks.* PLoS One, 2014. **9**(2): p. e86587.
134. Duarte, N.C., et al., *Global reconstruction of the human metabolic network based on genomic and bibliomic data.* Proceedings of the National Academy of Sciences, 2007. **104**(6): p. 1777.
135. Thiele, I., et al., *A community-driven global reconstruction of human metabolism.* Nat Biotechnol, 2013. **31**(5): p. 419-25.
136. Thiele, I. and B.O. Palsson, *A protocol for generating a high-quality genome-scale metabolic reconstruction.* Nat Protoc, 2010. **5**(1): p. 93-121.
137. Schomburg, I., et al., *BRENDA, the enzyme database: updates and major new developments.* Nucleic Acids Res, 2004. **32**(Database issue): p. D431-3.
138. Garcia Sanchez, C.E. and R.G. Torres Saez, *Comparison and analysis of objective functions in flux balance analysis.* Biotechnol Prog, 2014. **30**(5): p. 985-91.
139. Feist, A.M. and B.O. Palsson, *The biomass objective function.* Curr Opin Microbiol, 2010. **13**(3): p. 344-9.

140. Adadi, R., et al., *Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters.* PLoS Comput Biol, 2012. **8**(7): p. e1002575.
141. Folger, O., et al., *Predicting selective drug targets in cancer through metabolic networks.* Mol Syst Biol, 2011. **7**: p. 501.
142. Ramakrishna, R., et al., *Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints.* Am J Physiol Regul Integr Comp Physiol, 2001. **280**(3): p. R695-704.
143. Savinell, J.M. and B.O. Palsson, *Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism.* J Theor Biol, 1992. **154**(4): p. 421-54.
144. Fan, J., et al., *Quantitative flux analysis reveals folate-dependent NADPH production.* Nature, 2014. **510**(7504): p. 298-302.
145. Lee, D., et al., *Improving metabolic flux predictions using absolute gene expression data.* BMC Syst Biol, 2012. **6**: p. 73.
146. Montezano, D., et al., *Flux Balance Analysis with Objective Function Defined by Proteomics Data-Metabolism of Mycobacterium tuberculosis Exposed to Mefloquine.* PLoS One, 2015. **10**(7): p. e0134014.
147. Åkesson, M., J. Förster, and J. Nielsen, *Integration of gene expression data into genome-scale metabolic models.* Metabolic engineering, 2004. **6**(4): p. 285-293.
148. Becker, S.A. and B.O. Palsson, *Context-specific metabolic networks are consistent with experiments.* PLoS Comput Biol, 2008. **4**(5): p. e1000082.
149. Shlomi, T., et al., *Network-based prediction of human tissue-specific metabolism.* Nat Biotechnol, 2008. **26**(9): p. 1003-10.
150. Colijn, C., et al., *Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production.* PLoS Comput Biol, 2009. **5**(8): p. e1000489.
151. Chandrasekaran, S. and N.D. Price, *Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis.* Proceedings of the National Academy of Sciences, 2010. **107**(41): p. 17845-17850.
152. Jensen, P.A. and J.A. Papin, *Functional integration of a metabolic network model and expression data without arbitrary thresholding.* Bioinformatics, 2011. **27**(4): p. 541-7.
153. Agren, R., et al., *Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT.* PLoS computational biology, 2012. **8**(5).
154. Hoppe, A., S. Hoffmann, and H.-G. Holzhütter, *Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks.* BMC systems biology, 2007. **1**(1): p. 23.
155. Yizhak, K., et al., *Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model.* Bioinformatics, 2010. **26**(12): p. i255-i260.

156. Mahadevan, R., J.S. Edwards, and F.J. Doyle III, *Dynamic flux balance analysis of diauxic growth in Escherichia coli.* Biophysical journal, 2002. **83**(3): p. 1331-1340.
157. White, H.L. and J.R. White, *Lethal action and metabolic effects of streptonigrin on Escherichia coli.* Mol Pharmacol, 1968. **4**(6): p. 549-65.
158. Kappus, H. and H. Sies, *Toxic drug effects associated with oxygen metabolism: redox cycling and lipid peroxidation.* Experientia, 1981. **37**(12): p. 1233-41.
159. Pink, J.J., et al., *NAD(P)H:Quinone oxidoreductase activity is the principal determinant of beta-lapachone cytotoxicity.* J Biol Chem, 2000. **275**(8): p. 5416-24.
160. Huang, X., et al., *An NQO1 substrate with potent antitumor activity that selectively kills by PARP1-induced programmed necrosis.* Cancer Res, 2012. **72**(12): p. 3038-47.
161. Bey, E.A., et al., *An NQO1- and PARP-1-mediated cell death pathway induced in non-small-cell lung cancer cells by beta-lapachone.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11832-7.
162. Bey, E.A., et al., *Catalase abrogates beta-lapachone-induced PARP1 hyperactivation-directed programmed necrosis in NQO1-positive breast cancers.* Mol Cancer Ther, 2013. **12**(10): p. 2110-20.
163. Li, L.S., et al., *Modulating endogenous NQO1 levels identifies key regulatory mechanisms of action of beta-lapachone for pancreatic cancer therapy.* Clin Cancer Res, 2011. **17**(2): p. 275-85.
164. Cosentino, C., D. Grieco, and V. Costanzo, *ATM activates the pentose phosphate pathway promoting anti-oxidant defence and DNA repair.* Embo j, 2011. **30**(3): p. 546-55.
165. Boothman, D.A., S. Greer, and A.B. Pardee, *Potentiation of halogenated pyrimidine radiosensitizers in human carcinoma cells by beta-lapachone (3,4-dihydro-2,2-dimethyl-2H-naphtho[1,2-b]pyran- 5,6-dione), a novel DNA repair inhibitor.* Cancer Res, 1987. **47**(20): p. 5361-6.
166. Boothman, D.A., D.K. Trask, and A.B. Pardee, *Inhibition of potentially lethal DNA damage repair in human tumor cells by beta-lapachone, an activator of topoisomerase I.* Cancer Res, 1989. **49**(3): p. 605-12.
167. Li, L.S., et al., *NQO1-Mediated Tumor-Selective Lethality and Radiosensitization for Head and Neck Cancer.* Mol Cancer Ther, 2016. **15**(7): p. 1757-67.
168. Chakrabarti, G., D.E. Gerber, and D.A. Boothman, *Expanding antitumor therapeutic windows by targeting cancer-specific nicotinamide adenine dinucleotide phosphate-biogenesis pathways.* Clin Pharmacol, 2015. **7**: p. 57-68.
169. Chakrabarti, G., et al., *Tumor-selective use of DNA base excision repair inhibition in pancreatic cancer using the NQO1 bioactivatable drug, beta-lapachone.* Sci Rep, 2015. **5**: p. 17066.
170. Moore, Z., et al., *NAMPT inhibition sensitizes pancreatic adenocarcinoma cells to tumor-selective, PAR-independent metabolic catastrophe and cell death induced by beta-lapachone.* Cell Death Dis, 2015. **6**: p. e1599.
171. Chen, X., et al., *Analysis of DNA methylation and gene expression in radiation-resistant head and neck tumors.* Epigenetics, 2015. **10**(6): p. 545-61.

172. Fletcher, E.V., et al., *EGFR inhibition induces proinflammatory cytokines via NOX4 in HNSCC.* Mol Cancer Res, 2013. **11**(12): p. 1574-84.
173. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-27.
174. Wilhelm, M., et al., *Mass-spectrometry-based draft of the human proteome.* Nature, 2014. **509**(7502): p. 582-7.
175. Bansal, N., et al., *Broad phenotypic changes associated with gain of radiation resistance in head and neck squamous cell cancer.* Antioxid Redox Signal, 2014. **21**(2): p. 221-36.
176. Rebhan, M., et al., *GeneCards: integrating information about genes, proteins and diseases.* Trends Genet, 1997. **13**(4): p. 163.
177. Park, J.O., et al., *Metabolite concentrations, fluxes and free energies imply efficient enzyme usage.* Nat Chem Biol, 2016. **12**(7): p. 482-9.
178. Wishart, D.S., et al., *HMDB: the Human Metabolome Database.* Nucleic Acids Res, 2007. **35**(Database issue): p. D521-6.
179. Mims, J., et al., *Energy metabolism in a matched model of radiation resistance for head and neck squamous cell cancer.* Radiat Res, 2015. **183**(3): p. 291-304.
180. Chin, D., et al., *Novel markers for poor prognosis in head and neck cancer.* Int J Cancer, 2005. **113**(5): p. 789-97.
181. Cerna, D., et al., *Inhibition of nicotinamide phosphoribosyltransferase (NAMPT) activity by small molecule GMX1778 regulates reactive oxygen species (ROS)-mediated cytotoxicity in a p53-and nicotinic acid phosphoribosyltransferase1 (NAPRT1)-dependent manner.* Journal of Biological Chemistry, 2012. **287**(26): p. 22408-22417.
182. La Verde, N., et al., *Safe chemotherapy and hormone therapy for treating early breast cancer in a glucose 6-phosphate dehydrogenase-deficient patient: case report.* Anticancer Drugs, 2012. **23**(7): p. 758-60.
183. Lee, S.W., N. Chaiyakunapruk, and N.M. Lai, *What G6PD-deficient individuals should really avoid.* Br J Clin Pharmacol, 2017. **83**(1): p. 211-212.
184. Christmas, P., et al., *Alternative splicing determines the function of CYP4F3 by switching substrate specificity.* J Biol Chem, 2001. **276**(41): p. 38166-72.
185. Ross, D. and D. Siegel, *NAD(P)H:quinone oxidoreductase 1 (NQO1, DT-diaphorase), functions and pharmacogenetics.* Methods Enzymol, 2004. **382**: p. 115-44.
186. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
187. Wermuth, B., *Purification and properties of an NADPH-dependent carbonyl reductase from human brain. Relationship to prostaglandin 9-ketoreductase and xenobiotic ketone reductase.* J Biol Chem, 1981. **256**(3): p. 1206-13.
188. Dolfi, S.C., et al., *The metabolic demands of cancer cells are coupled to their size and protein synthesis rates.* Cancer Metab, 2013. **1**(1): p. 20.
189. Yizhak, K., et al., *Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer.* Elife, 2014. **3**.
190. Zielinski, D.C., et al., *Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism.* Sci Rep, 2017. **7**: p. 41241.

191. Mukherjee, P.K., et al., *Metabolomic analysis identifies differentially produced oral metabolites, including the oncometabolite 2-hydroxyglutarate, in patients with head and neck squamous cell carcinoma.* BBA Clin, 2017. **7**: p. 8-15.
192. Dinkova-Kostova, A.T. and P. Talalay, *NAD(P)H:quinone acceptor oxidoreductase 1 (NQO1), a multifunctional antioxidant enzyme and exceptionally versatile cytoprotector.* Arch Biochem Biophys, 2010. **501**(1): p. 116-23.
193. Smith, L., et al., *Proteomic identification of putative biomarkers of radiotherapy resistance: a possible role for the 26S proteasome?* Neoplasia, 2009. **11**(11): p. 1194-207.
194. Lee, Y.S., et al., *Differential gene expression profiles of radioresistant non-small-cell lung cancer cell lines established by fractionated irradiation: tumor protein p53-inducible protein 3 confers sensitivity to ionizing radiation.* Int J Radiat Oncol Biol Phys, 2010. **77**(3): p. 858-66.
195. Skvortsova, I., et al., *Intracellular signaling pathways regulating radioresistance of human prostate carcinoma cells.* Proteomics, 2008. **8**(21): p. 4521-33.
196. Brady, L.W., C.A. Perez, and D.E. Wazer, *Perez & Brady's principles and practice of radiation oncology*. 2013: Lippincott Williams & Wilkins.
197. Cadet, J. and J.R. Wagner, *DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation.* Cold Spring Harb Perspect Biol, 2013. **5**(2).
198. Tominaga, H., et al., *Involvement of reactive oxygen species (ROS) in the induction of genetic instability by radiation.* J Radiat Res, 2004. **45**(2): p. 181-8.
199. Harris, I.S., et al., *Glutathione and thioredoxin antioxidant pathways synergize to drive cancer initiation and progression.* Cancer Cell, 2015. **27**(2): p. 211-22.
200. Forshaw, T.E., et al., *Peroxiredoxins in Cancer and Response to Radiation Therapies.* Antioxidants (Basel), 2019. **8**(1).
201. Turgeon, M.O., N.J.S. Perry, and G. Poulogiannis, *DNA Damage, Repair, and Cancer Metabolism.* Front Oncol, 2018. **8**: p. 15.
202. Chatterjee, A., *Reduced glutathione: a radioprotector or a modulator of DNA-repair activity?* Nutrients, 2013. **5**(2): p. 525-42.
203. Alvarez-Idaboy, J.R. and A. Galano, *On the chemical repair of DNA radicals by glutathione: hydrogen vs electron transfer.* J Phys Chem B, 2012. **116**(31): p. 9316-25.
204. Schultz, A. and A.A. Qutub, *Reconstruction of Tissue-Specific Metabolic Networks Using CORDA.* PLoS Comput Biol, 2016. **12**(3): p. e1004808.
205. Zhang, C. and Q. Hua, *Applications of Genome-Scale Metabolic Models in Biotechnology and Systems Medicine.* Front Physiol, 2015. **6**: p. 413.
206. Lewis, J.E., et al., *Genome-Scale Modeling of NADPH-Driven beta-Lapachone Sensitization in Head and Neck Squamous Cell Carcinoma.* Antioxid Redox Signal, 2018. **29**(10): p. 937-952.
207. Lewis, J.E., et al., *Targeting NAD(+) Metabolism to Enhance Radiation Therapy Responses.* Semin Radiat Oncol, 2019. **29**(1): p. 6-15.
208. Kilburn, D.G., M.D. Lilly, and F.C. Webb, *The energetics of mammalian cell growth.* J Cell Sci, 1969. **4**(3): p. 645-54.

209. Fritzemeier, C.J., et al., *Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal.* PLoS Comput Biol, 2017. **13**(4): p. e1005494.
210. Kirkman, H.N., S. Galiano, and G.F. Gaetani, *The function of catalase-bound NADPH.* J Biol Chem, 1987. **262**(2): p. 660-6.
211. Rahman, M., et al., *Alternative preprocessing of RNA-Sequencing data in The Cancer Genome Atlas leads to improved analysis results.* Bioinformatics, 2015. **31**(22): p. 3666-72.
212. Schwanhausser, B., et al., *Global quantification of mammalian gene expression control.* Nature, 2011. **473**(7347): p. 337-42.
213. Wang, M., et al., *Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines.* Proteomics, 2015. **15**(18): p. 3163-8.
214. Binder, J.X., et al., *COMPARTMENTS: unification and visualization of protein subcellular localization evidence.* Database (Oxford), 2014. **2014**: p. bau012.
215. Dang, L., et al., *Cancer-associated IDH1 mutations produce 2-hydroxyglutarate.* Nature, 2009. **462**(7274): p. 739-44.
216. Avellaneda Matteo, D., et al., *Molecular mechanisms of isocitrate dehydrogenase 1 (IDH1) mutations identified in tumors: The role of size and hydrophobicity at residue 132 on catalytic efficiency.* J Biol Chem, 2017. **292**(19): p. 7971-7983.
217. Schellenberger, J., N.E. Lewis, and B.O. Palsson, *Elimination of thermodynamically infeasible loops in steady-state metabolic models.* Biophys J, 2011. **100**(3): p. 544-553.
218. Noor, E., et al., *Consistent estimation of Gibbs energy using component contributions.* PLoS Comput Biol, 2013. **9**(7): p. e1003098.
219. Price, P.J. and E.A. Gregory, *Relationship between in vitro growth promotion and biophysical and biochemical properties of the serum supplement.* In Vitro, 1982. **18**(6): p. 576-84.
220. Huang, X., et al., *Leveraging an NQO1 Bioactivatable Drug for Tumor-Selective Use of Poly(ADP-ribose) Polymerase Inhibitors.* Cancer Cell, 2016. **30**(6): p. 940-952.
221. Rahman, I., A. Kode, and S.K. Biswas, *Assay for quantitative determination of glutathione and glutathione disulfide levels using enzymatic recycling method.* Nat Protoc, 2006. **1**(6): p. 3159-65.
222. Gray, V.E., et al., *Quantitative Missense Variant Effect Prediction Using Large-Scale Mutagenesis Data.* Cell Syst, 2018. **6**(1): p. 116-124.e3.
223. Jones, M.E., *Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis.* Annu Rev Biochem, 1980. **49**: p. 253-79.
224. Schnell, J.R., H.J. Dyson, and P.E. Wright, *Structure, dynamics, and catalytic function of dihydrofolate reductase.* Annu Rev Biophys Biomol Struct, 2004. **33**: p. 119-40.
225. Kim, J. and R.J. DeBerardinis, *Mechanisms and Implications of Metabolic Heterogeneity in Cancer.* Cell Metab, 2019. **30**(3): p. 434-446.
226. Agarwal, A.R., F. Yin, and E. Cadenas, *Short-term cigarette smoke exposure leads to metabolic alterations in lung alveolar cells.* Am J Respir Cell Mol Biol, 2014. **51**(2): p. 284-93.

227. Noronha-Dutra, A.A., M.M. Epperlein, and N. Woolf, *Effect of cigarette smoking on cultured human endothelial cells.* Cardiovasc Res, 1993. **27**(5): p. 774-8.
228. Choi, Y.M., et al., *Mechanism of Cisplatin-Induced Cytotoxicity Is Correlated to Impaired Metabolism Due to Mitochondrial ROS Generation.* PLoS One, 2015. **10**(8): p. e0135083.
229. Chen, X., et al., *Modulators of Redox Metabolism in Head and Neck Cancer.* Antioxid Redox Signal, 2018. **29**(16): p. 1660-1690.
230. Cohen, A.L., S.L. Holmen, and H. Colman, *IDH1 and IDH2 mutations in gliomas.* Curr Neurol Neurosci Rep, 2013. **13**(5): p. 345.
231. Han, X., et al., *A molecular view of the radioresistance of gliomas.* Oncotarget, 2017. **8**(59): p. 100931-100941.
232. Molenaar, R.J., et al., *Radioprotection of IDH1-Mutated Cancer Cells by the IDH1-Mutant Inhibitor AGI-5198.* Cancer Res, 2015. **75**(22): p. 4790-802.
233. Lu, V.M. and K.L. McDonald, *Isocitrate dehydrogenase 1 mutation subtypes at site 132 and their translational potential in glioma.* CNS Oncol, 2018. **7**(1): p. 41-50.
234. Moreira, J.D., et al., *The Redox Status of Cancer Cells Supports Mechanisms behind the Warburg Effect.* Metabolites, 2016. **6**(4).
235. Zhuang, H., et al., *Downregulation of glycine decarboxylase enhanced cofilin-mediated migration in hepatocellular carcinoma cells.* Free Radic Biol Med, 2018. **120**: p. 1-12.
236. Bankar, S.B., et al., *Glucose oxidase--an overview.* Biotechnol Adv, 2009. **27**(4): p. 489-501.
237. Adimora, N.J., D.P. Jones, and M.L. Kemp, *A model of redox kinetics implicates the thiol proteome in cellular hydrogen peroxide responses.* Antioxid Redox Signal, 2010. **13**(6): p. 731-43.
238. Spitz, D.R., et al., *Metabolic oxidation/reduction reactions and cellular responses to ionizing radiation: a unifying concept in stress response biology.* Cancer Metastasis Rev, 2004. **23**(3-4): p. 311-22.
239. Tuttle, S.W., et al., *Sensitivity to chemical oxidants and radiation in CHO cell lines deficient in oxidative pentose cycle activity.* Int J Radiat Oncol Biol Phys, 1992. **22**(4): p. 671-5.
240. Varnes, M.E., *Inhibition of pentose cycle of A549 cells by 6-aminonicotinamide: consequences for aerobic and hypoxic radiation response and for radiosensitizer action.* NCI Monogr, 1988(6): p. 199-203.
241. Sattler, U.G., et al., *Glycolytic metabolism and tumour response to fractionated irradiation.* Radiother Oncol, 2010. **94**(1): p. 102-9.
242. Newman, W.C. and E.A. Monaco, 3rd, *Sensitization of Glioblastoma Cells to Irradiation by Modulating the Glucose Metabolism.* Neurosurgery, 2015. **77**(4): p. N16.
243. Meng, M.B., et al., *Targeting pyruvate kinase M2 contributes to radiosensitivity of non-small cell lung cancer cells in vitro and in vivo.* Cancer Lett, 2015. **356**(2 Pt B): p. 985-93.
244. Rashmi, R., et al., *Radioresistant Cervical Cancers Are Sensitive to Inhibition of Glycolysis and Redox Metabolism.* Cancer Res, 2018. **78**(6): p. 1392-1403.

245. Gudmundsson, S. and I. Thiele, *Computationally efficient flux variability analysis.* BMC Bioinformatics, 2010. **11**: p. 489.
246. Benfeitas, R., et al., *Characterization of heterogeneous redox responses in hepatocellular carcinoma patients using network analysis.* EBioMedicine, 2019. **40**: p. 471-487.
247. Porporato, P.E., et al., *Mitochondrial metabolism and cancer.* Cell Res, 2018. **28**(3): p. 265-280.
248. Ciccarese, F. and V. Ciminale, *Escaping Death: Mitochondrial Redox Homeostasis in Cancer Cells.* Front Oncol, 2017. **7**: p. 117.
249. Yin, F., H. Sancheti, and E. Cadenas, *Silencing of nicotinamide nucleotide transhydrogenase impairs cellular redox homeostasis and energy metabolism in PC12 cells.* Biochim Biophys Acta, 2012. **1817**(3): p. 401-9.
250. Jin, L., et al., *Glutamate dehydrogenase 1 signals through antioxidant glutathione peroxidase 1 to regulate redox homeostasis and tumor growth.* Cancer Cell, 2015. **27**(2): p. 257-70.
251. Stein, E.M., et al., *AG-221, an oral, selective, first-in-class, potent inhibitor of the IDH2 mutant metabolic enzyme, induces durable remissions in a phase I study in patients with IDH2 mutation positive advanced hematologic malignancies*. 2014, American Society of Hematology Washington, DC.
252. Hsieh, J.Y., et al., *A small-molecule inhibitor suppresses the tumor-associated mitochondrial NAD(P)+-dependent malic enzyme (ME2) and induces cellular senescence.* Oncotarget, 2015. **6**(24): p. 20084-98.
253. DeBerardinis, R.J. and N.S. Chandel, *Fundamentals of cancer metabolism.* Sci Adv, 2016. **2**(5): p. e1600200.
254. Follia, L., et al., *Integrative Analysis of Novel Metabolic Subtypes in Pancreatic Cancer Fosters New Prognostic Biomarkers.* Front Oncol, 2019. **9**: p. 115.
255. Stuani, L., M. Sabatier, and J.E. Sarry, *Exploiting metabolic vulnerabilities for personalized therapy in acute myeloid leukemia.* BMC Biol, 2019. **17**(1): p. 57.
256. Wallace, T.C., et al., *Personalized Nutrition in Disrupting Cancer - Proceedings From the 2017 American College of Nutrition Annual Meeting.* J Am Coll Nutr, 2019. **38**(1): p. 1-14.
257. Choi, Y.K. and K.G. Park, *Targeting Glutamine Metabolism for Cancer Treatment.* Biomol Ther (Seoul), 2018. **26**(1): p. 19-28.
258. Dahm-Daphi, J., C. Sass, and W. Alberti, *Comparison of biological effects of DNA damage induced by ionizing radiation and hydrogen peroxide in CHO cells.* Int J Radiat Biol, 2000. **76**(1): p. 67-75.
259. Erudaitius, D., et al., *Calculated cell-specific intracellular hydrogen peroxide concentration: Relevance in cancer cell susceptibility during ascorbate therapy.* Free Radic Biol Med, 2018. **120**: p. 356-367.
260. Adolfsen, K.J. and M.P. Brynildsen, *A Kinetic Platform to Determine the Fate of Hydrogen Peroxide in Escherichia coli.* PLoS Comput Biol, 2015. **11**(11): p. e1004562.
261. Vetrano, A.M., et al., *Characterization of the oxidase activity in mammalian catalase.* J Biol Chem, 2005. **280**(42): p. 35372-81.

262. Glorieux, C. and P.B. Calderon, *Catalase, a remarkable enzyme: targeting the oldest antioxidant enzyme to find a new cancer treatment approach.* Biol Chem, 2017. **398**(10): p. 1095-1108.
263. Middelkoop, E., et al., *Topology of catalase assembly in human skin fibroblasts.* Biochim Biophys Acta, 1993. **1220**(1): p. 15-20.
264. Lazarow, P.B. and C. de Duve, *The synthesis and turnover of rat liver peroxisomes. V. Intracellular pathway of catalase synthesis.* J Cell Biol, 1973. **59**(2 Pt 1): p. 507-24.
265. Radi, R., et al., *Detection of catalase in rat heart mitochondria.* J Biol Chem, 1991. **266**(32): p. 22028-34.
266. Bai, J. and A.I. Cederbaum, *Mitochondrial catalase and oxidative injury.* Biol Signals Recept, 2001. **10**(3-4): p. 189-99.
267. Bai, J., et al., *Overexpression of catalase in cytosolic or mitochondrial compartment protects HepG2 cells against oxidative injury.* J Biol Chem, 1999. **274**(37): p. 26217-24.
268. Kaushik, A.K. and R.J. DeBerardinis, *Applications of metabolomics to study cancer metabolism.* Biochim Biophys Acta Rev Cancer, 2018. **1870**(1): p. 2-14.
269. Willers, H., et al., *Basic mechanisms of therapeutic resistance to radiation and chemotherapy in lung cancer.* Cancer J, 2013. **19**(3): p. 200-7.
270. Jayasurya, K., et al., *Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy.* Med Phys, 2010. **37**(4): p. 1401-7.
271. Deist, T.M., et al., *Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers.* Med Phys, 2018. **45**(7): p. 3449-3459.
272. Kang, J., et al., *Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective.* Int J Radiat Oncol Biol Phys, 2015. **93**(5): p. 1127-35.
273. de Jong, M.C., et al., *Pretreatment microRNA Expression Impacting on Epithelial-to-Mesenchymal Transition Predicts Intrinsic Radiosensitivity in Head and Neck Cancer Cell Lines and Patients.* Clin Cancer Res, 2015. **21**(24): p. 5630-8.
274. Weichselbaum, R.R., et al., *An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer.* Proc Natl Acad Sci U S A, 2008. **105**(47): p. 18490-5.
275. Kim, H.S., et al., *Identification of a radiosensitivity signature using integrative metaanalysis of published microarray data for NCI-60 cancer cells.* BMC Genomics, 2012. **13**: p. 348.
276. Abazeed, M.E., et al., *Integrative radiogenomic profiling of squamous cell lung cancer.* Cancer Res, 2013. **73**(20): p. 6289-98.
277. Speers, C., et al., *Development and Validation of a Novel Radiosensitivity Signature in Human Breast Cancer.* Clin Cancer Res, 2015. **21**(16): p. 3667-77.
278. Piening, B.D., et al., *A radiation-derived gene expression signature predicts clinical outcome for breast cancer patients.* Radiat Res, 2009. **171**(2): p. 141-54.

279. Pitroda, S.P., et al., *DNA repair pathway gene expression score correlates with repair proficiency and tumor sensitivity to chemotherapy.* Sci Transl Med, 2014. **6**(229): p. 229ra42.
280. Tang, Z., et al., *Predicting radiotherapy response for patients with soft tissue sarcoma by developing a molecular signature.* Oncol Rep, 2017. **38**(5): p. 2814-2824.
281. Torres-Roca, J.F., et al., *Prediction of radiation sensitivity using a gene expression classifier.* Cancer Res, 2005. **65**(16): p. 7169-76.
282. Amundson, S.A., et al., *Integrating global gene expression and radiation survival parameters across the 60 cell lines of the National Cancer Institute Anticancer Drug Screen.* Cancer Res, 2008. **68**(2): p. 415-24.
283. Nuyten, D.S., et al., *Predicting a local recurrence after breast-conserving therapy by gene expression profiling.* Breast Cancer Res, 2006. **8**(5): p. R62.
284. Ghaffari, P., A. Mardinoglu, and J. Nielsen, *Cancer Metabolism: A Modeling Perspective.* Front Physiol, 2015. **6**: p. 382.
285. Zampieri, G., et al., *Machine and deep learning meet genome-scale metabolic modeling.* PLoS Comput Biol, 2019. **15**(7): p. e1007084.
286. Bionetworks, S. *Synapse.* [cited 2019 Sep 16]; Available from: https://www.synapse.org/#!Synapse:syn300013/wiki/27406.
287. Spainhour, J.C.G., J. Lim, and P. Qiu, *GDISC: a web portal for integrative analysis of gene-drug interaction for survival in cancer.* Bioinformatics, 2017. **33**(9): p. 1426-1428.
288. Cibulskis, K., et al., *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.* Nat Biotechnol, 2013. **31**(3): p. 213-9.
289. Chen, T. and C. Guestrin. *Xgboost: A scalable tree boosting system*. in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.
290. Bergstra, J., et al., *Hyperopt: a python library for model selection and hyperparameter optimization.* Computational Science & Discovery, 2015. **8**(1): p. 014008.
291. Lundberg, S.M. and S.-I. Lee. *A unified approach to interpreting model predictions*. in *Advances in neural information processing systems*. 2017.
292. Lundberg, S.M., et al., *Explainable ai for trees: From local explanations to global understanding.* arXiv preprint arXiv:1905.04610, 2019.
293. Yard, B.D., et al., *A genetic basis for the variation in the vulnerability of cancer to DNA damage.* Nat Commun, 2016. **7**: p. 11428.
294. Friedman, J.H., *Greedy function approximation: a gradient boosting machine.* Annals of statistics, 2001: p. 1189-1232.
295. Saeys, Y., I. Inza, and P. Larranaga, *A review of feature selection techniques in bioinformatics.* Bioinformatics, 2007. **23**(19): p. 2507-17.
296. Snoek, J., H. Larochelle, and R.P. Adams. *Practical bayesian optimization of machine learning algorithms*. in *Advances in neural information processing systems*. 2012.
297. Ghashghaei, M., et al., *Identification of a Radiosensitivity Molecular Signature Induced by Enzalutamide in Hormone-sensitive and Hormone-resistant Prostate Cancer Cells.* Sci Rep, 2019. **9**(1): p. 8838.

298. Hino, S., et al., *Cytoplasmic TSC-22 (transforming growth factor-beta-stimulated clone-22) markedly enhances the radiation sensitivity of salivary gland cancer cells.* Biochem Biophys Res Commun, 2002. **292**(4): p. 957-63.
299. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
300. Luo, J., et al., *mRNA and methylation profiling of radioresistant esophageal cancer cells: the involvement of Sall2 in acquired aggressive phenotypes.* J Cancer, 2017. **8**(4): p. 646-656.
301. Gong, L., et al., *Differential radiation response between normal astrocytes and glioma cells revealed by comparative transcriptome analysis.* Onco Targets Ther, 2017. **10**: p. 5755-5764.
302. Dahan, P., et al., *Ionizing radiations sustain glioblastoma cell dedifferentiation to a stem-like phenotype through survivin: possible involvement in radioresistance.* Cell Death Dis, 2014. **5**: p. e1543.
303. Heddleston, J.M., et al., *Hypoxia inducible factors in cancer stem cells.* Br J Cancer, 2010. **102**(5): p. 789-95.
304. Niu, N., et al., *Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines.* Genome Res, 2010. **20**(11): p. 1482-92.
305. Deng, Q., et al., *Chemotherapy and radiotherapy downregulate the activity and expression of DNA methyltransferase and enhance Bcl-2/E1B-19-kDa interacting protein-3-induced apoptosis in human colorectal cancer cells.* Chemotherapy, 2012. **58**(6): p. 445-53.
306. Hurov, K.E., C. Cotta-Ramusino, and S.J. Elledge, *A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability.* Genes Dev, 2010. **24**(17): p. 1939-50.
307. Barretina, J., et al., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.* Nature, 2012. **483**(7391): p. 603-7.
308. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
309. Kiefer, J., et al., *Abstract 3589: A systematic approach toward gene annotation of the hallmarks of cancer.* Cancer Research, 2017. **77**(13 Supplement): p. 3589-3589.
310. Lomax, M.E., L.K. Folkes, and P. O'Neill, *Biological consequences of radiation-induced DNA damage: relevance to radiotherapy.* Clin Oncol (R Coll Radiol), 2013. **25**(10): p. 578-85.
311. Morgan, W.F., et al., *Genomic instability induced by ionizing radiation.* Radiation research, 1996. **146**(3): p. 247-258.
312. Powell, S. and T.J. McMillan, *DNA damage and repair following treatment with ionizing radiation.* Radiother Oncol, 1990. **19**(2): p. 95-108.
313. Willers, H., J. Dahm-Daphi, and S.N. Powell, *Repair of radiation damage to DNA.* Br J Cancer, 2004. **90**(7): p. 1297-301.
314. Segal, E., et al., *A module map showing conditional activity of expression modules in cancer.* Nat Genet, 2004. **36**(10): p. 1090-8.
315. Poole, L.B., *The basics of thiols and cysteines in redox biology and chemistry.* Free Radic Biol Med, 2015. **80**: p. 148-57.

316. Forman, H.J., H. Zhang, and A. Rinna, *Glutathione: overview of its protective roles, measurement, and biosynthesis.* Mol Aspects Med, 2009. **30**(1-2): p. 1-12.
317. Zhao, L., et al., *Lipoamide Acts as an Indirect Antioxidant by Simultaneously Stimulating Mitochondrial Biogenesis and Phase II Antioxidant Enzyme Systems in ARPE-19 Cells.* PLoS One, 2015. **10**(6): p. e0128502.
318. Hou, Y., et al., *Lipoamide Ameliorates Oxidative Stress via Induction of Nrf2/ARE Signaling Pathway in PC12 Cells.* J Agric Food Chem, 2019. **67**(29): p. 8227-8234.
319. Persson, H.L., A.I. Svensson, and U.T. Brunk, *Alpha-lipoic acid and alpha-lipoamide prevent oxidant-induced lysosomal rupture and apoptosis.* Redox Rep, 2001. **6**(5): p. 327-34.
320. McBean, G.J., *Cysteine, Glutathione, and Thiol Redox Balance in Astrocytes.* Antioxidants (Basel), 2017. **6**(3).
321. Yurkova, I., et al., *Radiation-induced free-radical transformation of phospholipids: MALDI-TOF MS study.* Chem Phys Lipids, 2004. **132**(2): p. 235-46.
322. Laiakis, E.C., et al., *Metabolic phenotyping reveals a lipid mediator response to ionizing radiation.* J Proteome Res, 2014. **13**(9): p. 4143-54.
323. Barrera, G., *Oxidative stress and lipid peroxidation products in cancer progression and therapy.* ISRN Oncol, 2012. **2012**: p. 137289.
324. Institute, N.C. *Developmental Therapeutics Program (DTP) of the National Cancer Institute (NCI)*. [cited 2019 Dec 25]; Available from: https://wiki.nci.nih.gov/display/NCIDTPdata/Molecular+Target+Data.
325. Werner, E., et al., *Ionizing Radiation induction of cholesterol biosynthesis in Lung tissue.* Sci Rep, 2019. **9**(1): p. 12546.
326. Parasassi, T., et al., *Cholesterol protects the phospholipid bilayer from oxidative damage.* Free Radic Biol Med, 1995. **19**(4): p. 511-6.
327. Allalunis-Turner, M.J., et al., *Isolation of two cell lines from a human malignant glioma specimen differing in sensitivity to radiation and chemotherapeutic drugs.* Radiat Res, 1993. **134**(3): p. 349-54.
328. Huang, M.Y., et al., *CDC25A, VAV1, TP73, BRCA1 and ZAP70 gene overexpression correlates with radiation response in colorectal cancer.* Oncol Rep, 2011. **25**(5): p. 1297-306.
329. Moding, E.J., M.B. Kastan, and D.G. Kirsch, *Strategies for optimizing the response of cancer and normal tissues to radiation.* Nat Rev Drug Discov, 2013. **12**(7): p. 526-42.
330. Rhomberg, W., *The radiation response of sarcomas by histologic subtypes: a review with special emphasis given to results achieved with razoxane.* Sarcoma, 2006. **2006**(1): p. 87367.
331. Conklin, K.A., *Chemotherapy-associated oxidative stress: impact on chemotherapeutic effectiveness.* Integr Cancer Ther, 2004. **3**(4): p. 294-300.
332. Behera, R.K. and R. Nayak, *Expression profiling of nucleotide metabolism-related genes in human breast cancer cells after treatment with 5-fluorouracil.* Cancer Invest, 2009. **27**(5): p. 561-7.

333. Lawrence, T.S., A.W. Blackstock, and C. McGinn, *The mechanism of action of radiosensitization of conventional chemotherapeutic agents.* Semin Radiat Oncol, 2003. **13**(1): p. 13-21.
334. Revannasiddaiah, S. and S.P. Susheela, *Chemically enhanced radiotherapy: visions for the future.* Ann Transl Med, 2016. **4**(3): p. 52.
335. Kang, H.B., et al., *Metabolic Rewiring by Oncogenic BRAF V600E Links Ketogenesis Pathway to BRAF-MEK1 Signaling.* Mol Cell, 2015. **59**(3): p. 345-358.
336. Kim, S., et al., *Radiation-induced autophagy potentiates immunotherapy of cancer via up-regulation of mannose 6-phosphate receptor on tumor cells in mice.* Cancer Immunol Immunother, 2014. **63**(10): p. 1009-21.
337. Jaillet, C., et al., *Radiation-induced changes in the glycome of endothelial cells with functional consequences.* Sci Rep, 2017. **7**(1): p. 5290.
338. Lee, S.I. and K.S. Kang, *Function of capric acid in cyclophosphamide-induced intestinal inflammation, oxidative stress, and barrier function in pigs.* Sci Rep, 2017. **7**(1): p. 16530.
339. Kumar, A.P., et al., *Effect of butyric acid supplementation on serum and renal antioxidant enzyme activities in streptozotocin-induced diabetic rats.* Journal of food biochemistry, 2010. **34**: p. 15-30.
340. Gavino, V.C., et al., *Effect of polyunsaturated fatty acids and antioxidants on lipid peroxidation in tissue cultures.* J Lipid Res, 1981. **22**(5): p. 763-9.
341. Drouin, G., et al., *Comparative effects of dietary n-3 docosapentaenoic acid (DPA), DHA and EPA on plasma lipid parameters, oxidative status and fatty acid tissue composition.* J Nutr Biochem, 2019. **63**: p. 186-196.
342. Cameron, N.E. and M.A. Cotter, *Interaction between oxidative stress and gamma-linolenic acid in impaired neurovascular function of diabetic rats.* Am J Physiol, 1996. **271**(3 Pt 1): p. E471-6.
343. Henry, G.E., et al., *Antioxidant and cyclooxygenase activities of fatty acids found in food.* J Agric Food Chem, 2002. **50**(8): p. 2231-4.
344. Beckham, T.H., et al., *Interdiction of sphingolipid metabolism to improve standard cancer therapies.* Adv Cancer Res, 2013. **117**: p. 1-36.
345. Lewis, A.C., et al., *Targeting sphingolipid metabolism as an approach for combination therapies in haematological malignancies.* Cell Death Discov, 2018. **4**: p. 4.
346. Wei, S., et al., *Radiation-induced changes in nucleotide metabolism of two colon cancer cell lines with different radiosensitivities.* Int J Radiat Biol, 1999. **75**(8): p. 1005-13.
347. Moffatt, B.A. and H. Ashihara, *Purine and pyrimidine nucleotide synthesis and metabolism.* Arabidopsis Book, 2002. **1**: p. e0018.
348. Kaplan, H.S., J.D. Earle, and F.L. Howsden, *THE ROLE OF PURINE AND PYRIMIDINE BASES AND THEIR ANALOGUES IN RADIATION SENSITIVITY.* J Cell Comp Physiol, 1964. **64**: p. Suppl 1:69-89.
349. Bhatt, A.N., et al., *Transient elevation of glycolysis confers radio-resistance by facilitating DNA repair in cells.* BMC Cancer, 2015. **15**: p. 335.

350. Lu, C.L., et al., *Tumor cells switch to mitochondrial oxidative phosphorylation under radiation via mTOR-mediated hexokinase II inhibition--a Warburg-reversing effect.* PLoS One, 2015. **10**(3): p. e0121046.
351. Reisz, J.A., et al., *Effects of ionizing radiation on biological molecules--mechanisms of damage and emerging methods of detection.* Antioxid Redox Signal, 2014. **21**(2): p. 260-92.
352. Maxim, L.D., R. Niebo, and M.J. Utell, *Screening tests: a review with examples.* Inhal Toxicol, 2014. **26**(13): p. 811-28.
353. Hutchinson, J. and L. Marignol, *Clinical Potential of Statins in Prostate Cancer Radiation Therapy.* Anticancer Res, 2017. **37**(10): p. 5363-5372.
354. Solanki, A.A. and S.L. Liauw, *Role of HMG-CoA reductase inhibitors with curative radiotherapy in men with prostate cancer.* Open Access J Urol, 2011. **3**: p. 95-104.
355. Chen, Y.A., et al., *Simvastatin Sensitizes Radioresistant Prostate Cancer Cells by Compromising DNA Double-Strand Break Repair.* Front Pharmacol, 2018. **9**: p. 600.
356. Efimova, E.V., et al., *HMG-CoA Reductase Inhibition Delays DNA Repair and Promotes Senescence After Tumor Irradiation.* Mol Cancer Ther, 2018. **17**(2): p. 407-418.
357. Rae, C., et al., *Inhibition of Fatty Acid Synthase Sensitizes Prostate Cancer Cells to Radiotherapy.* Radiat Res, 2015. **184**(5): p. 482-93.
358. Wang, D. and R.N. Dubois, *Eicosanoids and cancer.* Nat Rev Cancer, 2010. **10**(3): p. 181-93.
359. Das, U.N., *Radiation resistance, invasiveness and metastasis are inflammatory events that could be suppressed by lipoxin A4.* Prostaglandins Leukot Essent Fatty Acids, 2012. **86**(1-2): p. 3-11.
360. Kim, K.Y., et al., *The combined treatment of aspirin and radiation induces apoptosis by the regulation of bcl-2 and caspase-3 in human cervical cancer cell.* Cancer Lett, 2003. **189**(2): p. 157-66.
361. Choy, H. and L. Milas, *Enhancing radiotherapy with cyclooxygenase-2 enzyme inhibitors: a rational advance?* J Natl Cancer Inst, 2003. **95**(19): p. 1440-52.
362. Gash, K.J., et al., *Potentiating the effects of radiotherapy in rectal cancer: the role of aspirin, statins and metformin as adjuncts to therapy.* Br J Cancer, 2017. **117**(2): p. 210-219.
363. Jacobs, C.D., et al., *Aspirin improves outcome in high risk prostate cancer patients treated with radiation therapy.* Cancer Biol Ther, 2014. **15**(6): p. 699-706.
364. Wolpert, D.H., *Stacked generalization.* Neural networks, 1992. **5**(2): p. 241-259.
365. Ma, Z., et al., *Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose.* PLoS One, 2018. **13**(10): p. e0205872.
366. Bhatt, S., et al., *Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization.* J R Soc Interface, 2017. **14**(134).
367. Grenet, I., et al., *Stacked Generalization with Applicability Domain Outperforms Simple QSAR on in Vitro Toxicological Data.* J Chem Inf Model, 2019. **59**(4): p. 1486-1496.

368. Ramalingam, P., M. Sandhya, and S. Sankar, *Using an innovative stacked ensemble algorithm for the accurate prediction of preterm birth.* J Turk Ger Gynecol Assoc, 2019. **20**(2): p. 70-78.
369. He, L., et al., *Extracting drug-drug interaction from the biomedical literature using a stacked generalization-based approach.* PLoS One, 2013. **8**(6): p. e65814.
370. Mirza, B., et al., *Machine Learning and Integrative Analysis of Biomedical Big Data.* Genes (Basel), 2019. **10**(2).
371. Kelly, C.J., et al., *Key challenges for delivering clinical impact with artificial intelligence.* BMC Med, 2019. **17**(1): p. 195.
372. Hanash, S.M., C.S. Baik, and O. Kallioniemi, *Emerging molecular biomarkers—blood-based strategies to detect and monitor cancer.* Nature reviews Clinical oncology, 2011. **8**(3): p. 142.
373. Gatenby, R.A., et al., *Adaptive therapy.* Cancer Res, 2009. **69**(11): p. 4894-903.
374. Yan, D. and D. Georg, *Adaptive radiation therapy.* Z Med Phys, 2018. **28**(3): p. 173-174.
375. Ree, A.H. and K.R. Redalen, *Personalized radiotherapy: concepts, biomarkers and trial design.* Br J Radiol, 2015. **88**(1051): p. 20150009.
376. Wardman, P., *Chemical radiosensitizers for use in radiotherapy.* Clin Oncol (R Coll Radiol), 2007. **19**(6): p. 397-417.
377. Luzhna, L., et al., *Molecular mechanisms of radiation resistance in doxorubicin-resistant breast adenocarcinoma cells.* Int J Oncol, 2013. **42**(5): p. 1692-708.
378. Jagetia, G.C. and V. Nayak, *Effect of doxorubicin on cell survival and micronuclei formation in HeLa cells exposed to different doses of gamma-radiation.* Strahlenther Onkol, 2000. **176**(9): p. 422-8.
379. Boeckman, H.J., K.S. Trego, and J.J. Turchi, *Cisplatin sensitizes cancer cells to ionizing radiation via inhibition of nonhomologous end joining.* Mol Cancer Res, 2005. **3**(5): p. 277-85.
380. El-Kareh, A.W. and T.W. Secomb, *A mathematical model for cisplatin cellular pharmacodynamics.* Neoplasia, 2003. **5**(2): p. 161-9.
381. Chen, T.L., et al., *Nonlinear pharmacokinetics of cyclophosphamide and 4-hydroxycyclophosphamide/aldophosphamide in patients with metastatic breast cancer receiving high-dose chemotherapy followed by autologous bone marrow transplantation.* Drug Metab Dispos, 1997. **25**(5): p. 544-51.
382. Sahoo, S., et al., *Modeling the effects of commonly used drugs on human metabolism.* Febs j, 2015. **282**(2): p. 297-317.
383. Bordel, S., *Constraint based modeling of metabolism allows finding metabolic cancer hallmarks and identifying personalized therapeutic windows.* Oncotarget, 2018. **9**(28): p. 19716-19729.
384. Garcia Sanchez, C.E., C.A. Vargas Garcia, and R.G. Torres Saez, *Predictive potential of flux balance analysis of Saccharomyces cerevisiae using as optimization function combinations of cell compartmental objectives.* PLoS One, 2012. **7**(8): p. e43006.
385. Budinich, M., et al., *A multi-objective constraint-based approach for modeling genome-scale microbial ecosystems.* PLoS One, 2017. **12**(2): p. e0171744.
386. Matsson, P., et al., *Quantifying the impact of transporters on cellular drug permeability.* Trends Pharmacol Sci, 2015. **36**(5): p. 255-62.

387. Eckford, P.D. and F.J. Sharom, *ABC efflux pump-based resistance to chemotherapy drugs.* Chem Rev, 2009. **109**(7): p. 2989-3011.
388. Dethmers, J.K. and A. Meister, *Glutathione export by human lymphoid cells: depletion of glutathione by inhibition of its synthesis decreases export and increases sensitivity to irradiation.* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7492-6.
389. Zhu, Y., et al., *Emerging evidence for targeting mitochondrial metabolic dysfunction in cancer therapy.* J Clin Invest, 2018. **128**(9): p. 3682-3691.
390. Ortega, A.L., S. Mena, and J.M. Estrela, *Glutathione in cancer cell death.* Cancers (Basel), 2011. **3**(1): p. 1285-310.
391. DeLong, E.R., D.M. DeLong, and D.L. Clarke-Pearson, *Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.* Biometrics, 1988. **44**(3): p. 837-45.
392. Raaphorst, G.P., M. Leblanc, and L.F. Li, *A comparison of response to cisplatin, radiation and combined treatment for cells deficient in recombination repair pathways.* Anticancer Res, 2005. **25**(1a): p. 53-8.
393. Wojcik, M., et al., *Glutamine-dependent NAD+ synthetase. How a two-domain, three-substrate enzyme avoids waste.* J Biol Chem, 2006. **281**(44): p. 33395-402.
394. Pollak, N., M. Niere, and M. Ziegler, *NAD kinase levels control the NADPH concentration in human cells.* J Biol Chem, 2007. **282**(46): p. 33562-71.
395. Koc, G., et al., *Does cisplatin-based chemotherapy effect on blood lipid levels of patients with germ cell testicular tumor in long-term follow-up?* Int Urol Nephrol, 2011. **43**(4): p. 1095-100.
396. Al-Bahlani, S., et al., *Fatty acid synthase regulates the chemosensitivity of breast cancer cells to cisplatin-induced apoptosis.* Apoptosis, 2017. **22**(6): p. 865-876.
397. Sheng, H., et al., *Omega-3 Polyunsaturated Fatty Acids Enhance Cisplatin Efficacy in Gastric Cancer Cells by Inducing Apoptosis via ADORA1.* Anticancer Agents Med Chem, 2016. **16**(9): p. 1085-92.
398. Marinello, A.J., et al., *Inhibition of NADPH-cytochrome P450 reductase by cyclophosphamide and its metabolites.* Biochem Biophys Res Commun, 1981. **99**(2): p. 399-406.
399. Liu, Y., H. Peng, and J.T. Zhang, *Expression profiling of ABC transporters in a drug-resistant breast cancer cell line using AmpArray.* Mol Pharmacol, 2005. **68**(2): p. 430-8.
400. Elliott, A.M. and M.A. Al-Hajj, *ABCB8 mediates doxorubicin resistance in melanoma cells by protecting the mitochondrial genome.* Mol Cancer Res, 2009. **7**(1): p. 79-87.
401. Konkimalla, V.B., B. Kaina, and T. Efferth, *Role of transporter genes in cisplatin resistance.* In Vivo, 2008. **22**(3): p. 279-83.
402. Bartoszek, A., *Metabolic activation of adriamycin by NADPH-cytochrome P450 reductase; overview of its biological and biochemical effects.* Acta Biochim Pol, 2002. **49**(2): p. 323-31.
403. Chen, H.H. and M.T. Kuo, *Role of glutathione in the regulation of Cisplatin resistance in cancer chemotherapy.* Met Based Drugs, 2010. **2010**.

404. Kasherman, Y., S. Sturup, and D. Gibson, *Is glutathione the major cellular target of cisplatin? A study of the interactions of cisplatin with cancer cell extracts.* J Med Chem, 2009. **52**(14): p. 4319-28.
405. Kato, T., et al., *Cisplatin and radiation sensitivity in human head and neck squamous carcinomas are independently modulated by glutathione and transcription factor NF-kappaB.* Head Neck, 2000. **22**(8): p. 748-59.
406. Lai, G.M., et al., *Effect of glutathione on DNA repair in cisplatin-resistant human ovarian cancer cell lines.* J Natl Cancer Inst, 1989. **81**(7): p. 535-9.
407. Taymaz-Nikerel, H., et al., *Doxorubicin induces an extensive transcriptional and metabolic rewiring in yeast cells.* Sci Rep, 2018. **8**(1): p. 13672.
408. Wu, K.C., J.Y. Cui, and C.D. Klaassen, *Beneficial role of Nrf2 in regulating NADPH generation and consumption.* Toxicol Sci, 2011. **123**(2): p. 590-600.
409. Ma, L. and Q. Cheng, *Inhibiting 6-phosphogluconate dehydrogenase reverses doxorubicin resistance in anaplastic thyroid cancer via inhibiting NADPH-dependent metabolic reprogramming.* Biochem Biophys Res Commun, 2018. **498**(4): p. 912-917.
410. Covert, M.W., et al., *Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli.* Bioinformatics, 2008. **24**(18): p. 2044-50.
411. Lee, J.M., et al., *Dynamic analysis of integrated signaling, metabolic, and regulatory networks.* PLoS Comput Biol, 2008. **4**(5): p. e1000086.
412. Jaramillo, M.C. and D.D. Zhang, *The emerging role of the Nrf2-Keap1 signaling pathway in cancer.* Genes Dev, 2013. **27**(20): p. 2179-91.
413. Kansanen, E., et al., *The Keap1-Nrf2 pathway: Mechanisms of activation and dysregulation in cancer.* Redox Biol, 2013. **1**: p. 45-9.
414. Lin, T.-Y., L.C. Cantley, and G.M. DeNicola, *NRF2 rewires cellular metabolism to support the antioxidant response.* A Master Regulator of Oxidative Stress-The Transcription Factor Nrf2, 2016.
415. Damiani, C., et al., *popFBA: tackling intratumour heterogeneity with Flux Balance Analysis.* Bioinformatics, 2017. **33**(14): p. i311-i318.
416. Ghadiri, M., et al., *A multiscale agent-based framework integrated with a constraint-based metabolic network model of cancer for simulating avascular tumor growth.* Mol Biosyst, 2017. **13**(9): p. 1888-1897.
417. Lin, J., et al., *The roles of glucose metabolic reprogramming in chemo- and radio-resistance.* J Exp Clin Cancer Res, 2019. **38**(1): p. 218.
418. Wagner, W., W.M. Ciszewski, and K.D. Kania, *L- and D-lactate enhance DNA repair and modulate the resistance of cervical carcinoma cells to anticancer drugs via histone deacetylase inhibition and hydroxycarboxylic acid receptor 1 activation.* Cell Commun Signal, 2015. **13**: p. 36.
419. Dimopoulos, A.C. and M. Reczko, *Biomass distances from personalized multispecies dynamic flux balance analysis of the human gut microbiome identify dietary influences for patients with and without inflammatory bowel disease.* 2019.
420. Cooper, L.A., et al., *PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective.* J Pathol, 2018. **244**(5): p. 512-524.

421. Mobadersany, P., et al., *Predicting cancer outcomes from histology and genomics using convolutional networks.* Proc Natl Acad Sci U S A, 2018. **115**(13): p. E2970-e2979.
422. Miousse, I.R., K.R. Kutanzi, and I. Koturbash, *Effects of ionizing radiation on DNA methylation: from experimental biology to clinical applications.* Int J Radiat Biol, 2017. **93**(5): p. 457-469.
423. Chaudhry, M.A. and R.A. Omaruddin, *Differential DNA methylation alterations in radiation-sensitive and -resistant cells.* DNA Cell Biol, 2012. **31**(6): p. 908-16.
424. Tomasik, B., W. Fendler, and D. Chowdhury, *Serum microRNAs - potent biomarkers for radiation biodosimetry.* Oncotarget, 2018. **9**(18): p. 14038-14039.
425. Czochor, J.R. and P.M. Glazer, *microRNAs in cancer cell response to ionizing radiation.* Antioxid Redox Signal, 2014. **21**(2): p. 293-312.
426. Yousefi, S., et al., *Learning Clinical Outcomes from Heterogeneous Genomic Data Sources.* arXiv preprint arXiv:1904.01637, 2019.
427. Sun, X. and B. Hu, *Mathematical modeling and computational prediction of cancer drug resistance.* Brief Bioinform, 2018. **19**(6): p. 1382-1399.
428. Kim, O.D., M. Rocha, and P. Maia, *A Review of Dynamic Modeling Approaches and Their Application in Computational Strain Optimization for Metabolic Engineering.* Front Microbiol, 2018. **9**: p. 1690.
429. Feng, X., et al., *Integrating flux balance analysis into kinetic models to decipher the dynamic metabolism of Shewanella oneidensis MR-1.* PLoS Comput Biol, 2012. **8**(2): p. e1002376.
430. Chowdhury, A., A.R. Zomorrodi, and C.D. Maranas, *k-OptForce: integrating kinetics with flux balance analysis for strain design.* PLoS Comput Biol, 2014. **10**(2): p. e1003487.